THE
PHYSICAL REVIEW
A Journal of Experimental and Theoretical Physics Established by E. L. Nichols in 1893

VOL. 55, NO. 8 APRIL 15, 1939 SECOND SERIES

Band Spectra in Nitrogen at Atmospheric Pressure. A Source of Band Spectra Excitation

OLIVER R. WULF AND EUGENE H. MELVIN
Bureau of Chemistry and Soils, U. S. Department of Agriculture, Washington, D. C.
(Received February 20, 1939)

An apparatus, essentially an ozonizer, consisting of a cylindrical condenser, a portion of the dielectric of which is a gas, has been used as a source of band spectra excitation. Such a source can be operated with the gas at atmospheric pressure and at a temperature which is practically that of the walls. Nitrogen bands $A \rightarrow X$ have been photographed in emission at atmospheric pressure under resolution sufficient to permit measurement of part of their rotational structure. The rotational structure of four of the present bands has been computed by using the rotational constants which have been determined from other band systems in nitrogen. The agreement with the measured values is satisfactory, and makes more certain the identification of this band system. The alternation of intensity in the rotational lines is clearly visible, measurements indicating that the ratio of the intensities is 2 : 1. In this source of excitation the first positive group of nitrogen is considerably weaker relative to the second positive group than in an ordinary glow discharge. The rather striking difference between this type of source and a high voltage, low current arc at atmospheric pressure is illustrated by band spectra in oxygen containing a small amount of nitrogen. In the latter source the principal emission is the Runge bands of oxygen while in the former the only prominent material is the second positive group of nitrogen.

AN ORDINARY all-glass ozonizer in operation is luminescent, and affords a source of band spectra excitation which may be of use because of the unique conditions under which it operates. In particular two conditions, which are at times desirable, can be met by using such a source; first, the gas may be excited at atmospheric pressure and second, it may, at the same time, be held at room temperature. Either one of these conditions can be met in other sources but ordinarily not both at the same time. In arc and flame excitation, while at atmospheric pressure, the gas is usually at high temperature while in the glow discharge the gas is at low pressure. In the present work emission spectra of nitrogen from the above-mentioned source have been studied, and, in particular, it has been found that certain bands otherwise difficult to excite may be obtained in this way.

APPARATUS

While there are many forms of ozonizers, the type used in the present work, shown in Fig. 1, consisted of two coaxial glass tubes separated by a gas space of a few millimeters. The outside of the outer cylinder and the inside of the inner cylinder were covered with metal coatings, thus forming essentially a Leyden jar with a portion of the dielectric a gas. The inlet and outlet tubes at either end were connected by a closed circuit of glass tubing consisting of a trap for immersion in liquid air or a portion of tubing containing

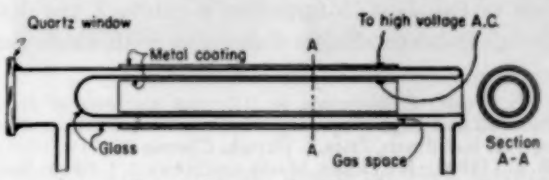


FIG. 1. Source of excitation.

material such as metallic sodium to be used as a purifying agent. A quartz window was cemented on the end of the ozonizer, through which the emission could be viewed. The spectrograph was so oriented that the extended plane of the slit and the collimater axis was tangent to the cylindrical annular, luminous gas space.

Tank nitrogen dried in liquid air was used in the work. The gamma-bands of NO, arising from the small amount of oxygen in the nitrogen, showed strongly in the emission, and since these bands interfered with the observation of the bands of interest, several methods were tried in an effort to remove the oxygen. First, a tungsten filament was heated in the gas. This formed large amounts of tungsten oxide without appreciably reducing the intensity of the NO bands. Second, hydrogen was added with a liquid-air trap in the system to freeze out the water formed. In order to reduce greatly the intensity of the NO bands, an excess of hydrogen was required. This brought out a continuous background of emission. Third, metallic sodium was introduced in the system and distilled along it, reducing greatly the intensity of the NO bands but not removing them completely. Fourth, aluminum foil was placed in the discharge space. This reduced the intensity of the NO bands to about the same extent as did the sodium. It appears that the excitation of the NO bands is a very sensitive test for small amounts of oxygen in nitrogen. All plates in the present work used for measurement were taken with an apparatus in which the oxygen had been reduced by means of metallic sodium. A large Hilger quartz spectrograph, with a 170-cm focus, was employed.

Mechanism of Excitation

The discharge in the ozonizer consists of a myriad of tiny sparks which may under certain conditions become so finely and uniformly distributed as to cause the appearance of a uniform glow. The character of the discharge is similar to that of the Tesla luminescence,[1] though the discharge between double dielectrics with small gas space separation is probably more completely distributed in the form of many fine discharges than is the Tesla luminescence as commonly used, and it can be produced without special apparatus.

The mechanism of the ozonizer action has been the subject of a large amount of work[2] which we shall not attempt to discuss here. The character of the discharge, however, seems to be in accord with the following description. Since an ozonizer does not operate under steady direct-current voltages, some important difference must be found in the conditions when alternating or pulsating direct voltages are applied. On the rise of voltage in the first quarter cycle there is a sweeping of the residual ions in the gas to the walls of the dielectric. Were the voltage change to stop here, that is, as in the application of a steady direct-current voltage, no further change would occur in the gas. In striking the walls of the dielectric, the ions, which are swept out of the gas by the rise of voltage in the first quarter cycle, eject electrons from the dielectric material. Thereafter these are also included in the process of sweeping, and they produce ionization in the gas as they move under the potential gradient. If the frequency is sufficiently high, as is the case at 60 cycles, this ionization produced in the gas does not disappear entirely in the time between voltage peaks. The ionization rises from the moment of applying the alternating potential until it comes to a steady-state condition, at which the average rate of the removal of ions by sweeping and by recombination is equal to the average rate at which ions are being produced.

Thus a short time after the application of the alternating potential the gas has become a conducting medium, still perhaps of rather high resistance, but nevertheless able to give appreciable currents at the voltages applied. Discharge phenomena occur in the gas space, therefore, but these are manifestations not so much of breakdown as of the carrying of current by the partially ionized gas. Now it is known that in a low pressure gas discharge, the cross section of the discharge decreases with increasing pressure. At

[1] N. Tesla, *Experiments in Alternate Currents of High Potential and High Frequency* (McGraw-Hill, 1904), p. 29 *et seq.*; H. Kauffman, Zeits. f. physik. Chemie **26**, 719 (1898); **28**, 673 (1899); McVicker, Marsh and Stewart, J. Chem. Soc. **123**, 642 (1923); **129**, 17 (1926).

[2] See for example G. Glockler and S. C. Lind, *The Electrochemistry of Gases and Other Dielectrics* (John Wiley and Sons, New York); M. Moeller, *Das Ozon*, (Vieweg, Braunschweig, 1921).

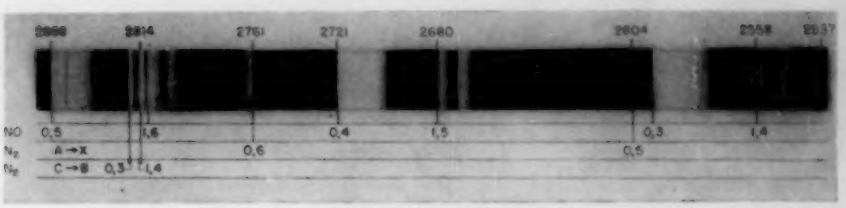


FIG. 2. Portion of ultraviolet emission from ozonizer filled with nitrogen at atmospheric pressure.

a given pressure, moreover, the cross section of the current stream in the gas decreases with decreasing current. Thus with increase of gas pressure and decrease of the current carried, the cross section of the current stream decreases. In the source of excitation used in the present work, one has these two conditions in the extreme; the pressure is atmospheric, and, because no single onset of current can discharge more than a small area of the ion layer on the dielectric surface, the current is very low. This results in a localized discharge which is probably of very small cross section. Such discharges occur at many places along the surface of the dielectric since, because no single discharge can clear more than a very small area, another such discharge can occur in the immediate neighborhood. Each of the tiny sparks consists of an extremely weak electric discharge. The total effect represents, nevertheless, appreciable light, especially when the annular space of the ozonizer is observed end-on. When the apparatus shown in Fig. 1 is operated at 15,000–25,000 volts and 60 cycles, the current indicated by an ordinary a.c. milliameter is of the order of a milliampere.

RESULTS AND DISCUSSION

With dry nitrogen at atmospheric pressure in the ozonizer, the emission observed consisted principally of the second positive group of nitrogen and the gamma-bands of NO, which appear to be of roughly the same intensity. The first positive group of nitrogen in this source is very much weaker relative to the second positive group than in an ordinary glow discharge. Other spectra can be seen on the plates and in particular the bands which have been observed by Vegard[3] and by Kaplan[4] and which are believed to be due to the intersystem transition $A\,^3\Sigma_u^+ \rightarrow X\,^1\Sigma_g^+$. Their intensity is relatively weak and the emission of the NO bands interferes greatly in observing them. However, by reducing the oxygen content a few of these intersystem transition bands were observed free of overlapping emission.

Figure 2 is an enlargement of a portion of a spectrogram showing some of the intersystem nitrogen bands in which part of the rotational structure is clearly visible. The rotational structure, although distinct, cannot, however, be accurately measured because of low resolution; this is due partly to the wide slit employed (0.05 mm), but chiefly to variation in room temperature during the exposure times of seven to eight hours. In the vicinity of the heads the lines are not resolved. In one weaker exposure with a slit width of 0.02 mm and less than the ordinary temperature variation, the lines were sharper and the diffuse lines near the head were weakened relative to those further removed. This indicated that they were composed of multiple overlapping lines. The rotational structure has been partially measured for the bands 0−5, 0−6, 1−4 and 1−5 and is given in Table I, the most satisfactory measurements are those for the bands 0−5 and 0−6.

The rotational constants for both the upper and lower levels of the $A \rightarrow X$ transition have been determined from other band systems in nitrogen better than could be done from the

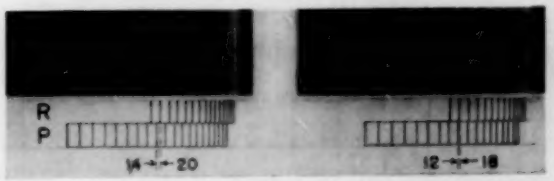


FIG. 3. The bands 0−6 and 0−5 of the $A \rightarrow X$ transition in N_2.

[3] L. Vegard, Zeits. f. Physik **75**, 30 (1932).
[4] J. Kaplan, Phys. Rev. **44**, 947 (1933); **45**, 675 (1934); and **45**, 898 (1934). Nature **141**, 645 (1938) and **141**, 1139 (1938).

FIG. 4. Comparison of ozonizer and arc excitation in oxygen.

present spectrograms. In Table I the observed values for the lines are compared with those computed from the rotational constants given by Sponer[5] and from an estimated position of the origin of each band. Each band is considered to consist of a single P and R branch, and further fine structure is neglected. The agreements seem to be as good as can be expected and make more certain the identification of the emitter of the bands. As shown in Table I higher members of the R branch lie in nearly the same positions as lower members of the P branch. In Table II a comparison is made of three sets of combination differences. It seems probable that, because of the relative intensities, the measured values almost correspond to the P branch lines. Fig. 3 shows the bands $0-5$ and $0-6$ with a portion of the rotational structure of each sketched below. The alternation of intensity in the band lines can be clearly seen.

It happens that it is the stronger lines of each branch which nearly overlap so that the intensity alternation is visible, and is not partially masked as would be the case if the stronger lines of the one branch overlapped the weaker lines of the other branch. An approximate determination of the intensity ratio, based on ten lines from one plate in each of the two bands $0-5$ and $0-6$, gave 1.94 as the average of sixteen values. These lay in the range 1.82 to 2.08 and were individually the average of the two ratios for three consecutive lines. This latter averaging was made to correct for the general change in intensity of the lines with change of quantum number. Correction was not made for background. The result indicates that the ratio is 2 : 1, in accord with the two units of spin ($i=1$) possessed by the N^{14} nucleus.

TABLE I. *Rotational structure of* $(0-6)$, $(0-5)$, $(1-5)$ *and* $(1-4)$ *bands.*

K''	P Calc. (cm^{-1}) vac.	Obs. (cm^{-1}) vac.	R Calc. (cm^{-1}) vac.	K''
2761A band $(0-6)$ estimated origin at 36201.0 cm^{-1}				
		*	36203.9	0
		*	205.8	1
		*	206.9	2
		36207.5	207.1	3
		*	206.4	4
		*	204.8	5
		*	202.3	6
1	36197.2	*	198.9	7
2	192.6	193.6	194.6	8
3	187.0	*	189.5	9
4	180.6	182.3	183.4	10
5	173.2	*	176.4	11
6	165.0	166.0	168.5	12
7	155.9	156.0	159.8	13
8	145.9	147.1	150.1	14
9	135.0	134.0	139.6	15
10	123.1	123.7	128.2	16
11	110.4	110.2	115.8	17
12	96.9	99.1	102.6	18
13	82.4	83.2	88.5	19
14	67.0	69.2	73.5	20
15	50.7	52.2	57.6	21
16	33.5	36.3	40.8	22
17	15.5	16.5		
18	35996.5	35998.6		
19	976.6	979.3		
20	955.9	958.3		
21	934.3	938.5		
22	911.7	916.7		
2604A band $(0-5)$ estimated origin at 38389.0 cm^{-1}				
		*	38391.9	0
		*	393.8	1
		*	394.8	2
		*	394.9	3
	38394.7		394.0	4
		*	392.3	5
		*	389.5	6
1	38385.2	*	385.9	7
2	380.5	380.4	381.3	8
3	374.8	*	375.8	9
4	368.2	368.3	369.4	10
5	360.7	359.4	362.0	11
6	352.3	352.2	353.7	12
7	342.9	345.3	344.4	13
8	332.6	334.1	334.4	14
9	321.3	321.8	323.3	15
10	309.2	310.3	311.3	16
11	296.1	297.5	298.3	17
12	282.0	283.0	284.5	18
13	267.1	267.8	269.7	19
14	251.2	251.8	253.9	20
15	234.4	235.4		
16	216.6	218.1		
17	198.0	199.6		
18	178.3	180.0		
19	157.8	156.8		
20	136.3	139.5		
2510A band $(1-5)$ estimated origin at 39820.0 cm^{-1}				
		*	39822.8	0
		*	824.7	1
		*	825.7	2
		*	825.6	3
		39823.6	824.6	4
		*	822.7	5
		*	819.8	6
1	39816.2	*	815.9	7
2	811.4	811.8	811.1	8
3	805.7	*	805.3	9
4	799.1	800.0	798.6	10
5	791.4	*	790.9	11
6	782.9	782.9	782.3	12
7	773.3	774.2	772.7	13
8	762.8	763.0	762.1	14
9	751.4	752.2	750.6	15
10	738.9	738.9	738.1	16
11	725.6	726.1	724.7	17
12	711.3	710.5	710.3	18
13	696.0	699.3		
14	679.7	680.8		
15	662.5	665.5		
16	644.4	645.5		
17	625.3	626.4		
18	605.2	606.5		
2378A band $(1-4)$ estimated origin at 42036.0 cm^{-1}				
		*	42038.8	0
		*	40.7	1
	42041.5		41.5	2
		*	41.4	3
		*	40.3	4
		*	38.2	5
		*	35.0	6
1	42032.2	*	30.9	7
2	27.3	25.3	25.8	8
3	21.5	*	19.7	9
4	14.7	13.2	12.6	10
5	6.9	*	4.6	11
6	41998.1	41997.4	41995.5	12
7	988.3	987.0	985.4	13
8	977.5	976.2	974.4	14
9	965.7	964.6	962.3	15
10	953.0	952.5		
11	939.2	938.8		
12	925.5	924.8		
13	907.7	907.8		
14	892.0	893.2		
15	874.2			
16	855.5	853.6		

* The first measured lines are broad and diffuse and cover the starred assignments.

TABLE II. *Combination differences of the bands.*

K''	$R(K)-P(K)$ $(0-6)$	$(0-5)$	Calc.	$R(K-1)-P(K+1)$ $(0-5)$	$(1-5)$	Calc.	$R(K)-P(K)$ $(1-5)$	$(1-4)$	Calc.
12	66.9	69.2	71.7	91.6		94.9	72.4	72.6	71.0
13	72.8	77.5	77.3	100.4	102.1	102.5	74.9	79.2	76.7
14	77.9	82.3	83.2	109.9	108.7	110.0	82.2	83.0	82.4
15	81.8	86.4	88.9	116.0	117.5	117.8	86.7		88.1
16	87.4	92.2	94.7	122.2	125.8	125.3	93.4	98.9	93.7
17	93.7	97.9	100.3	130.3	132.4	133.0	99.7		99.4
18	100.5	103.0	106.2	140.7		140.5	104.0		105.1
19	103.9	111.0	111.9	143.5		148.2			
20	110.9	112.3	116.9						

[5] H. Sponer, *Molekülspektren* (Springer, Berlin, 1935).

The source of excitation used in the present work possesses some usefulness in band spectra research because of the two conditions of high gas pressure and low temperature that can be simultaneously met. The difference between two sources of excitation which do not have these two factors in common is illustrated in Fig. 4 in which some emission of oxygen containing a small amount of nitrogen is shown. The emission in the upper picture is from the ozonizer while in the lower picture it is from a high voltage, low current arc operated by the same type of transformer. In both cases the gas was at atmospheric pressure. In the lower picture the principal emission is the Runge bands of oxygen,[6] while from the ozonizer it is the second positive group of nitrogen. A further study is being made of the mechanism of excitation in the ozonizer.

The authors gratefully acknowledge the assistance given them by Mrs. L. S. Deming and Mr. R. T. O'Connor in the course of this work.

[6] C. Runge, Physica 1, 254 (1921).

APRIL 15, 1939 PHYSICAL REVIEW VOLUME 55

The Analysis of Nuclear Binding Energies

WALTER H. BARKAS
Institute for Advanced Study, Princeton, New Jersey
(Received February 25, 1939)

The fine structure of the mass defect curve has been studied, especially with reference to the nuclear symmetry character. From the analysis it has been found possible to obtain satisfactory empirical curves for the functions L and E_0 appearing in Wigner's theory. The behavior of these curves allows deductions to be made regarding nuclear shells, and binding energies of both known and as yet unknown, unstable, nuclei. About 150 masses have been computed in this way.

INTRODUCTION

THE light nuclei are well known to exhibit a marked periodicity of four, which may be interpreted as a consequence of the operation of the Pauli principle in systems containing two kinds of heavy particles. One may obtain an instructive representation of this behavior by plotting, as a function of the mass number A, the change in binding energy observed on the addition of successive particles. Such a plot (corresponding to the familiar graphs of ionization potentials for atomic electrons) is shown in Fig. 1. The effect is exhibited for two methods of building up atoms:

$$\nu+\mathrm{H}+\nu+\mathrm{H}+\nu+\mathrm{H}+\cdots,$$

and

$$\nu+\nu+\mathrm{H}+\mathrm{H}+\nu+\nu+\cdots,$$

where ν and H represent, respectively, a neutron and a hydrogen atom. The following facts are made evident without a detailed study by this method of displaying the empirical information on atomic masses.

(1) Successive particles entering the same period or level are bound with energies increasing in a roughly linear way for small mass numbers, but later a depression of the proton points takes place.

(2) The heights of the peaks decrease as the mass number increases.

(3) The binding energy of the last proton in Ne^{20} is abnormally high compared to that in O^{16}. This is probably to be correlated with the unusually low total binding energy of Ne^{20} noticed by Hafstad and Teller.[1] (We shall later adduce evidence that Ne^{20} is at the end of the $2s$ sub-shell.)

E. Wigner[2] has studied the consequences of a detailed application of the Pauli principle to a many-body nuclear model in which the specifically nuclear forces are equal between all pairs of particles and do not depend on the spin ("1th approximation"). Although it is now well known, particularly from the neutron-proton scattering

[1] L. R. Hafstad and E. Teller, Phys. Rev. 54, 684 (1938).
[2] E. Wigner, Phys. Rev. 51, 947 (1937).

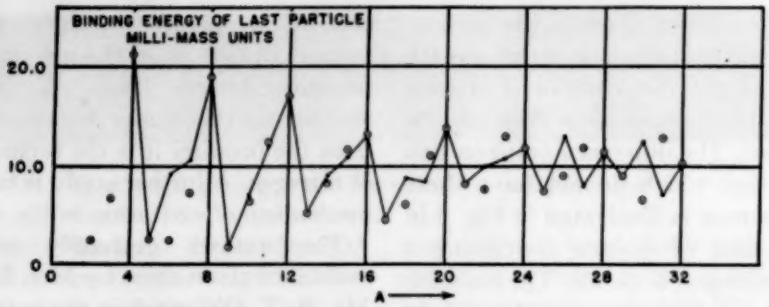


FIG. 1. Binding energy of successive nuclear particles as a function of the mass number. The connected points are determined from the series $p+p+n+n+p+p+\cdots$. The encircled points come from the sequence $p+n+p+n+p+\cdots$. The points affected by the mass of He^5 are uncertain.

experiments, and from the reported[3] quadrupole moment of deuterium, that spin dependent forces do actually exist, it seems that in most nuclei they play only a subordinate role so far as the binding energy is concerned.

The binding energy $-E$ was found to consist, under these assumptions, of the sum of four terms as follows:

$$-E=-E_0(A)-\Xi'L(A)+\tfrac{1}{2}K_1A^{-1}(A-T_t-1)T_t-K_2T_t^2/A. \quad (1)$$

These terms, in this order, are generally decreasing in magnitude. The first, $-E_0$, is a smoothly increasing function of A. The factor Ξ' is a number which represents the symmetry character of the configuration, while L is again a smooth function of A, decreasing as A^{-1} for large A. $T_t=\frac{1}{2}(N-Z)$ is the isotopic spin of the nucleus, and K_1 is a constant determined by the Coulomb energy under the assumption that the nuclear volume is proportional to A. The last term, which gives the dependence of kinetic energy on isotopic spin, while perhaps not strictly applicable to actual nuclei, is the best estimate available. The last two terms for many light nuclei are to be regarded as hardly more than small corrections. In this paper we shall undertake to examine the experimental findings by comparing them with this formula. The central problem will then be to investigate the behavior of the functions L and E_0.

Equation (1) may for the purpose of this study be written

$$-E_0-\Xi'L=-E-\tfrac{1}{2}K_1A^{-1}(A-T_t-1)T_t+K_2T_t^2/A, \quad (2)$$

where known or measurable terms appear on the right. The method of analysis will be to obtain as accurate estimates as possible for these terms, and then to graph the function $-E_0-\Xi'L$ for each value of Ξ'. (The value of Ξ' corresponding to a given nucleus provides a convenient and significant property for its classification.) Several curves are thus obtained, the ordinate separations of which determine L as a function of A. After one has the function L, $-E_0$ may be derived most accurately by using the $k\alpha$ type[4] nuclei, but where a rapid change of slope takes place other nuclear types have also been used to obtain points at other mass numbers. In the next section the various terms on the right side of Eq. (2) will be evaluated and examined as critically as present information permits.

NEGLECTED EFFECTS

The formula (1) is simple and unambiguous, so that when compared with the measured binding energies any important influences neglected in the derivation are detectable. As will be

[3] Kellogg, Rabi, Ramsey and Zacharias, Phys. Rev. 55, 318 (1939).

[4] A configuration $k\alpha$ (an alpha-particle type nucleus or one having $2k$ protons and $2k$ neutrons) may for convenience be referred to as a nuclear "core," because such a structure is known to bind individual protons and neutrons strongly. Other nuclei will be represented by $k\alpha$ plus additional neutrons and protons; e.g. $4\alpha+3p+n=F^{19}$. Such a "core" will only have the usual significance of the word if the additional particles are loosely bound, and if the "core" has no low lying excited states. These conditions are sometimes fulfilled in light nuclei. (O^{16} seems to have its first excited state at over 10 Mev.) In such cases the core will produce something very similar to a central field in which we may have some confidence that the orbital angular momentum will have properties similar to those studied extensively in connection with atomic spectra. The circumstance that spins of light nuclei may be predicted with some confidence suggests also that $L-S$ coupling is a valid assumption.

shown presently, for light nuclei it is feasible to correct for these. On the other hand, for heavy atoms this treatment is not adequate. Among these possible additional effects the following may be listed: (a) Additional spin dependent forces between the nuclear particles, (b) an effect on the symmetry character of the lowest state due to large Coulomb energy, (c) a strong tendency for shell or group structure, (d) a different dependence of kinetic energy on isotopic spin than that adopted, and (e) the electrostatic energy not so simply dependent on mass number and charge as supposed.

(a) When the mass values are studied it is observed that, if the assumption of a symmetrical interaction not depending on spin between all pairs of nuclear particles is made, the nuclei $(k\alpha+\nu+\pi)$ are more tightly bound compared to $(k\alpha+2\nu)$ than would be anticipated merely from the mass difference between the proton and neutron, after correction for the Coulomb energy. An accurate estimate of the electrostatic energy may be obtained from the isobaric pairs $(k\alpha+\nu)$, $(k\alpha+\pi)$, if the sole difference in binding energy for these isobars is this Coulomb energy.[5]

One is led to believe also that the kinetic energies are not different in the isobars $(k\alpha+2\nu)$, $(k\alpha+\pi+\nu)$ because the Pauli principle in each case permits both extra particles to enter the same level. (For nuclei with high Coulomb energy this situation may, however, be altered.) Hence any difference in the binding energies, corrected for Coulomb energy, appears to be attributable to a difference in the interaction between the "extra" particles. Since the spins of $(k\alpha+2\nu)$ nuclei are zero, and those of $(k\alpha+\pi+\nu)$ nuclei unity in every case investigated, we might designate these interactions by ${}^1\nu\nu$ and ${}^3\pi\nu$. That such a relatively large spin dependence exists has already been deduced from experiments on slow neutron scattering by hydrogen. Table I exhibits the present evidence for the effect.[6]

The Coulomb energies were derived from the inspection limits of beta-ray spectra when available. Otherwise they were computed by the usual method.[2] At mass numbers greater than 17 the Coulomb energy formula has not been adequately verified, so no great reliance should be placed on the differences at higher mass numbers. Indeed, the experiments of Pollard, Schultz and Brubaker,[7] and Brubaker[8] make it appear that considerable deviations from a constant ratio of nuclear volume to mass number may exist. Another interpretation of these experiments may be that the binding energy of a helium nucleus to A^{40} is particularly great because the addition of an alpha-particle to A^{40} carries one from a point somewhat off the bottom of the Heisenberg valley back to it. Thus the virtual levels of Ca^{44} may be reached with comparatively low α-particle energies, and this may be responsible for the anomalous penetration and scattering.

Although the spin dependent component of the nuclear forces is here shown to be sufficiently large to be detected definitely in the binding energies, it produces, after all, rather a small effect and will not cause any great difference in the behavior of actual nuclei as compared to the nuclear model based on a symmetric spin-independent Hamiltonian. The average of the

TABLE I. *Effect of spin dependent forces.*

Isobaric pair	Binding energy difference mMU	Coulomb energy difference mMU	2γ
He^6, Li^6	−3.2	1.4	4.6
Be^{10}, B^{10}	0.2	2.4	2.2
C^{14}, N^{14}	0.6	3.3	2.7
O^{18}, F^{18}	2.7	3.9	1.2
Ne^{22}, Na^{22}	2.5	4.6	2.1
Mg^{26}, Al^{26}	3.9	5.2	1.3
Si^{30}, P^{30}	5.1	5.7	0.6

[5] The assumption that such homologous nuclei have binding energies differing only by the Coulomb energy appears to depend just on the premise that the average specific interaction between pairs of protons in the nucleus is the same as the similar average interaction between neutrons. This is accepted as a principal assumption, the approximate validity of which may be established by a study of Fig. 1 or otherwise.

[6] The effect is expected to decrease with increasing nuclear radius as the average separation of the particles increases with the size of the nucleus. Even the irregularities seem to be real, however. A more complete study of this phenomenon seems to require an understanding of, or at least detailed assumptions regarding, the forces between individual elementary particles. Feenberg and Phillips (Phys. Rev. **51**, 597 (1937)) have observed also that possibly the order of levels in B^{10} and N^{14} is not that given by the Hartree model. Another matter causes difficulty. Not all the beta-activities for these isobars lie on the same Sargent curve, nor are the deviations systematic. For these reasons we avoid using the mass values of atoms of types $(k\alpha+2\nu)$ and $(k\alpha+\nu+\pi)$ in the determination of the functions L and E_0 where small discrepancies might produce spurious kinks in the curves.

[7] Pollard, Schultz and Brubaker, Phys. Rev. **53**, 351 (1938).

[8] G. Brubaker, Phys. Rev. **54**, 1011 (1938).

${}^3\pi\nu$ and the ${}^1\nu\nu$ interaction is approximately the average specific interaction between nuclear particles. If we write ${}^3\pi\nu - {}^1\nu\nu = 2\gamma$, then a reasonable procedure seems to be merely to correct for the effect in Eq. (1) by adding a term γ to the right-hand side of the equation for nuclei of type $(k\alpha+\nu+\pi)$, and subtracting it for nuclei $(k\alpha+2\nu)$ (and also presumably for $(k\alpha+2\pi)$).

(b) Disintegration energies are known very accurately for heavy, naturally radioactive atoms, and differences of binding energies are readily deduced. Consistent values for the constants appearing in the binding energy formula were not obtained, however, in many attempts at their determination. The most promising of these methods, that of utilizing the energy balance in the two branches of ThC, RaC and AcC proved no more successful than other devices. The conclusion to be drawn from this result is doubtless that the symmetry character of the ground state is affected by the large electrostatic energy, and that probably other terms, such as the Coulomb energy itself, are affected also.

(c) It might be expected on general grounds that the behavior of the function $-E_0(A)$ should yield information regarding possible "shells" in nuclei, for, if no such structures exist, the curve should be smoothly increasing with A. When the form of this function is obtained the evidence for such shells will be exhibited. On the other hand, an estimate of the amount of correlation between groups of particles such as those making up an α-particle is difficult. This is especially so, since according to Hafstad and Teller[1] no change of symmetry is introduced by postulating an α-particle structure. An α-particle model which assumes a correlation of *unity* between the positions of groups of four particles has very limited validity, however. This is demonstrated, for example, by the occurrence with a fairly large cross section of the type of reaction

$$k\alpha+\alpha \rightarrow (k+1)\alpha^* \rightarrow (k\alpha+2\nu+\pi)+\pi$$

when the energy balance permits. It seems that further information relative to the nature and probability of such grouping effects may be obtained from the several possible disintegration experiments in which the compound nucleus is of the $k\alpha$ type. α-particle projectiles are now available with sufficiently high energies so that almost all possible products of disintegration may be anticipated. It is probable, however, that the α-particle model is a better approximation for a nucleus in the ground state than in an excited state. The number of possible states of excitation should also depend on the amount of particle grouping.

(d) The expression $K_2T_t^2/A$ for the dependence of kinetic energy on T_t is for most light nuclei a small and doubtful term. The form of this expression erroneously implies a difference in kinetic energy for the nuclear types $(k\alpha+2\nu)$, $(k\alpha+\pi+\nu)$, which has already been discussed. This term will therefore be omitted for nuclei $(k\alpha+2\nu)$. For the larger values of T_t the value 26.4 mMU for K_2 will be used.[2] Another effect which occurs at the end of the $3d$ shell has been evaluated by Wigner.[2] Particles forced outside the closed shell appear to lose binding energy of approximately 1.6 mMU. A similar effect may be expected also at the end of other shells.

(e) The expression for the electrostatic energy of the last proton might be expected to vary with the radial extension of its wave function and therefore on its tightness of binding. A formula for this energy allowing for such an effect has been derived by Bethe.[9] For reasons of simplicity, however, this more refined formula will not be used; the error thus introduced is very slight. It will be shown next that a very satisfactory constant value is obtained for K_1 by using the formula for this constant derived from Eq. (1) on the lighter elements.*

TABLE II. *Determination of coefficient of Coulomb energy term.*

A	ΔE mMU	K_1
3	0.86	1.241
7	1.87	1.193
11	3.14	1.397
13	3.27	1.281
15	3.66	1.289
17	3.91	1.257
	Average	1.276

[9] H. A. Bethe, Phys. Rev. **54**, 436 (1938).

* *Note added in proof.*—As remarked earlier in this paper the Coulomb energy formula has not been tested adequately for $A>17$. This should perhaps be emphasized even more, particularly since the rather poorly known beta-disintegration energy of Si^{27} was not used in the evaluation of K_1 because it gave a value differing widely from the other numbers obtained, so that it was not considered wise to include this figure in averaging.

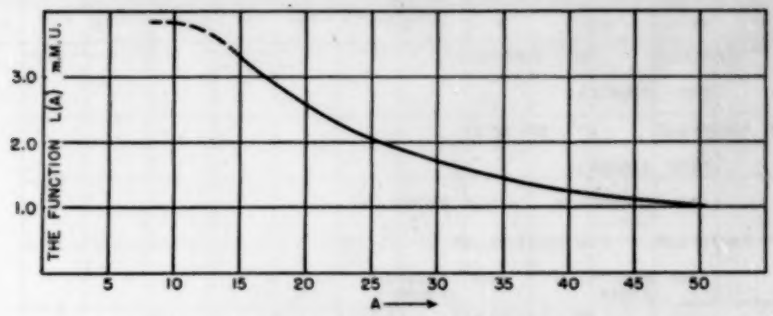


FIG. 2. The function $L(A)$ determined from the empirical masses.

From the isobaric pairs $(k\alpha+\nu)$, $(k\alpha+\pi)$, and $(k\alpha+\nu+2\pi)$, $(k\alpha+\pi+2\nu)$ the Coulomb energy may be determined, for on subtracting their binding energy equations, one is left with the expression

$$K_1=2A^{\frac{1}{3}}\Delta E/(A-1),$$

where ΔE is the binding energy difference for the isobars. The data from which K_1 is determined are given in Table II.

THE FUNCTIONS $L(A)$ AND $E_0(A)$

Figure 2 is a curve for $L(A)$. For $A>36$ very incomplete mass data are available, but in many cases L values may be obtained from the inspection upper limits of β-ray spectra when the absence of γ-rays has been established. The curve in this interval is to be regarded as considerably more tentative than for intermediate mass numbers. Likewise for very small mass numbers some approximations become poor. For example, it is difficult to be certain that the expression which has been used will give the correct kinetic energy difference between the isobars Li^8-Be^8, so that the behavior of the curve for small mass numbers may be misleading. Indeed, the values for L obtained by different methods do not check well below $A=12$; it is probable that the kinetic energy term, which here becomes quite large for certain nuclear types, is a very poor approximation in this region. For this reason the curve has not been carried below $A=8$. It is found, however, that almost the entire useful range of the curve $L(A)$ is represented rather accurately by the hyperbola $AL=50$, in agreement with theory.[9]

The values of Ξ' may be obtained according to the rules given in Wigner's paper.[9] Table III gives the numerical values taken on by Ξ' for some nuclear types.

It is anticipated that the curve for $L(A)$ may be found useful in the study of excited states.[10] It is notable in this connection that excited states of $k\alpha$ type nuclei which may be put in correspondence with the ground states of the isobaric nuclei $((k-1)\alpha+\pi+3\nu)$ do seem to exist. The present paper is, however, concerned principally with the normal configurations.

By using the curve derived for L one may add $\Xi'L$ to $-E_0-\Xi'L$ and obtain the curve for $-E_0$. This was done for $A=4k\pm1$ by using $(k\alpha+\nu)$, $(k\alpha+\pi)$, $(k\alpha+2\nu+\pi)$ and $(k\alpha+\nu+2\pi)$. Above $A=40$ the points were determined by using Sc^{45}, Ti^{48} and V^{51}.

The curve, Fig. 3, for E_0 makes a very definite dip in the region 16–20. Beyond 20 it rises smoothly to 40 where another break in the curve occurs. The points beyond 40 do not lie on a continuation of the previous section. Other circumstances likewise indicate that a change[11] takes place in the curve at this point.

TABLE III. *The ground state symmetry character for light nuclei.*

A	$T_\zeta=0$	$\pm1/2$	±1	$\pm3/2$	±2	$\pm5/2$	±3	$\pm7/2$	±4
$4k$	5/2		13/2		17/2		29/2		37/2
$4k+1$		35/8		59/8		91/8		131/8	
$4k+2$	5		5		25/2		13		20
$4k+3$		35/8		59/8		91/8		131/8	

[10] L. Motz and E. Feenberg, Phys. Rev. **54**, 1055 (1938); J. Bardeen and E. Feenberg, Phys. Rev. **54**, 809 (1938).

[11] It was remarked earlier that the change in kinetic energy taking place at $A=40$ can be detected. It may be noticed, also, that on a T_ζ *vs.* A diagram the locus of stable nuclei turns definitely to positive values of T_ζ beyond 40, while up to that point zero and even negative values were possible.

TABLE IV. *Computed and observed atomic masses. The experimental[12] value appears above the computed mass. Parentheses indicate a possible uncertainty ≦ 1 mMU. P.E. in 10^{-5} MU.*

A	$T_z=-1$	$-1/2$	0	$1/2$	1	$3/2$	2	$5/2$	3	$7/2$	4
1		↑ H^1 1.00812±2		↑ n^1 1.00893±5							
2			↑ H^2 2.01472±2								
3		↑ He^3 3.01701±12		↑ H^3 3.01704±7							
4			↑ He^4 4.00388±7								
6			↑ Li^6 6.01690±20		↑ He^6 6.02090±30						
7		↑ Be^7 7.01911±30		↑ Li^7 7.01804±20							
8			↑ Be^8 8.00777±20 (8.0078)		↑ Li^8 8.0249±30 (8.0250)						
9		↑ B^9 ..... (9.0164)		↑ Be^9 9.01497±25 (9.0151)		↑ Li^9 ..... (9.0313)					
10			↑ B^{10} 10.01605±30 (10.0144)		↑ Be^{10} 10.01671±30 (10.0148)						
11		↑ C^{11} 11.01544±35 11.0151		↑ B^{11} 11.01286±20 11.0130		↑ Be^{11} ..... (11.0277)					
12	↑ N^{12} 12.0233		↑ C^{12} 12.00398±10 12.0040		↑ B^{12} 12.0193±70 12.0188						
13		↑ N^{13} 13.01005±15 13.0101		↑ C^{13} 13.00766±15 13.0077		↑ B^{13} ..... (13.0207)					
14	↑ O^{14} ..... (14.0136)		↑ N^{14} 14.00750±8 (14.0080)		↑ C^{14} 14.00763±12 (14.0083)						
15		↑ O^{15} 15.0078±40 15.0080		↑ N^{15} 15.00489±20 15.0051		↑ C^{15} ..... (15.0165)					
16	↑ F^{16} ..... (16.0175)		↑ O^{16} 16.00000 16.0000		↑ N^{16} 16.011±200 16.0114						
17		↑ F^{17} 17.0076±30 17.0075		↑ O^{17} 17.00450±7 17.0044		↑ N^{17} ..... (17.0136)					
18	↑ Ne^{18} ..... (18.0114)		↑ F^{18} 18.0066±100 (18.0066)		↑ O^{18} 18.0047±100 (18.0047)						
19		↑ Ne^{19} 19.0077		↑ F^{19} 19.00452±17 19.0042		↑ O^{19} ..... (19.0139)					
20	↑ Na^{20} ..... (20.0160)		↑ Ne^{20} 19.99881±11 19.9988		↑ F^{20} 20.0087±100 (20.0066)		↑ O^{20} ..... (20.0168)				
21		↑ Na^{21} 21.0035		↑ Ne^{21} 20.99968±23 20.9996		↑ F^{21} ..... (21.0059)		↑ O^{21} ..... (21.0176)			
22	↑ Mg^{22} ..... (22.0062)		↑ Na^{22} 22.0002±50 (21.9999)		↑ Ne^{22} 21.99864±36 (21.9983)		↑ F^{22} ..... (22.0157)				
23		↑ Mg^{23} ..... (23.0004)		↑ Na^{23} 22.9961±35 22.9961		↑ Ne^{23} 23.0013		↑ F^{23} ..... (23.0114)			
24	↑ Al^{24} ..... (24.0059)		↑ Mg^{24} 23.9924±60 23.9920		↑ Na^{24} 23.9974±45 23.9974		↑ Ne^{24} ..... (24.0012)		↑ F^{24} ..... (24.0160)		
25		↑ Al^{25} 24.9970		↑ Mg^{25} 24.9938±90 24.9926		↑ Na^{25} ..... (24.9967)		↑ Ne^{25} ..... (25.0054)		↑ F^{25} ..... (25.0186)	
26	↑ Si^{26} ..... (25.9993)		↑ Al^{26} 25.9929±200 (25.9935)		↑ Mg^{26} 25.9898±50 (25.9905)		↑ Na^{26} ..... (26.0044)		↑ Ne^{26} ..... (26.0069)		
27		↑ Si^{27} 26.9931±150 26.9944		↑ Al^{27} 26.9899±80 26.9896		↑ Mg^{27} 26.9921±150 26.9929		↑ Na^{27} ..... (27.0003)		↑ Ne^{27} ..... (27.0119)	
28	↑ P^{28} 28.0001		↑ Si^{28} 27.9866±60 27.9870		↑ Al^{28} 27.9903±70 27.9904		↑ Mg^{28} ..... (27.9926)		↑ Na^{28} ..... (28.0041)		
29		↑ P^{29} 28.9919		↑ Si^{29} 28.9866±60 28.9869		↑ Al^{29} 28.9904±200 28.9892		↑ Mg^{29} ..... (28.9956)		↑ Na^{29} ..... (29.0060)	
30	↑ S^{30} ..... (29.9944)		↑ P^{30} 29.9882±150 (29.9891)		Si^{30} 29.9832±90 (29.9841)		↑ Al^{30} ..... (29.9954)		↑ Mg^{30} ..... (29.9964)		
31		↑ S^{31} 30.9892		↑ P^{31} 30.9842±50 30.9841		↑ Si^{31} 30.9862±60 30.9858		↑ Al^{31} ..... (30.9911)		↑ Mg^{31} ..... (31.0002)	
32	↑ Cl^{32} 31.9950		↑ S^{32} 31.9823±30 31.9823		↑ P^{32} 31.9841±50 31.9841		↑ Si^{32} 31.9849		↑ Al^{32} ..... (31.9939)		
33		↑ Cl^{33} 32.9872		↑ S^{33} 32.9816		↑ P^{33} 32.9826		↑ Si^{33} ..... (32.9870)		↑ Al^{33} ..... (32.9950)	
34	↑ A^{34} ..... (33.9898)		↑ Cl^{34} 33.981±300 (33.9839)		↑ S^{34} 33.978±200 (33.9784)		↑ P^{34} ..... (33.9874)		↑ Si^{34} ..... (33.9873)		
35		↑ A^{35} 34.9851		↑ Cl^{35} 34.9803±60 34.9792		↑ S^{35} 34.9796		↑ P^{35} ..... (34.9833)		↑ Si^{35} ..... (34.9920)	
36	↑ K^{36} 35.9907		↑ A^{36} 35.9780±100 35.9783		↑ Cl^{36} 35.978±100 35.9788		↑ S^{36} 35.978±100 35.9784		↑ P^{36} ..... (35.9870)		
37		↑ K^{37} 36.9830		↑ A^{37} 36.9769		↑ Cl^{37} 37.9779±120 36.9767		↑ S^{37} ..... (36.9812)		↑ P^{37} ..... (36.9889)	
38	↑ Ca^{38} ..... (37.9860)		↑ K^{38} ..... (37.9795)		↑ A^{38} 37.974±250 (37.9735)		↑ Cl^{38} 37.981±300 (37.9821)		↑ S^{38} ..... (37.9826)		
39		↑ Ca^{39} ..... (38.9811)		↑ K^{39} ..... (38.9747)		↑ A^{39} ..... (38.9755)		↑ Cl^{39} ..... (38.9794)			
40			↑ Ca^{40} 39.974±100 (39.9738)		↑ K^{40} 39.975±100 (39.9748)		↑ A^{40} 39.97504±26 (39.9750)		↑ Cl^{40} ..... (39.9819)		
45				↑ Ti^{45} ..... (44.9711)		↑ Sc^{45} 44.9689±100 44.9689		↑ Ca^{45} ..... (44.9693)			
48							↑ Ti^{48} 47.9651±50 47.9651				
51						↑ Cr^{51} 50.9613±100 (50.9611)		↑ V^{51} 50.9598±100 50.9598		↑ Ti^{51} 50.9612±100 (50.9610)	

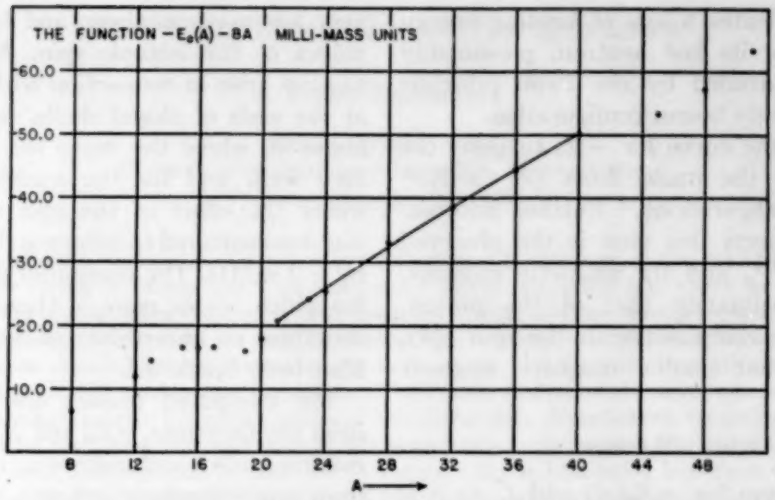


FIG. 3. $-E_0-8A$ as a function of the mass number. The term $8A$ is subtracted in order to reduce the ordinate values and bring out the structure of the curve. The function is expected to be smooth through the range of mass numbers where the $3d$ shell is being built up. Such a curve is drawn through this interval.

[13] References for experimental data: For the masses n^1, H^1, H^2, H^3, He^3, He^4, Li^6, Li^7, Be^8, Be^9, B^{10}, B^{11}, C^{12} and C^{13} I am indebted to Professor H. A. Bethe who kindly communicated this list to me. This combination at the moment seems most reliable. In arriving at these figures the disintegration data of Allison *et al.* (Phys. Rev. **54**, 171 (1938) and **55**, 107 (1939)) were utilized along with the best value for the binding energy of the deuteron. (H. A. Bethe, Phys. Rev. **53**, 314 (1938)). Other disintegration data used were $H^2(d, n)He^3$, 3.26 Mev and $Li^6(d, \alpha)He^4$, 22.21 Mev. The uncertainties given were estimated by the writer. With these data the mass difference $Be^{10}-B^{10}$ is, however, suspiciously large in view of the relative stability of Be^{10}.

He^6 comes from the present value of Li^6 and the beta-disintegration energy of Bjerge and Brostrom, Danske Videns. Selskab, Math. fys. Med. XVI, 8.

He^8 is not listed, for although Joliot and Zlotowski (J. de phys. et rad. **9**, 403 (1938)) report that it is stable with mass 5.0106 ± 0.0005, this is very difficult to interpret, for such a figure implies the instability of Be^9. By examining Fig. 1, one might be led to conclude that the stability of He^5 is improbable.

Li^8 is derived from the disintegration energy of Rumbaugh, Roberts and Hafstad (Phys. Rev. **54**, 657 (1938)) taken with the present value of Li^7.

Be^7 is derived from the energy limits determined from a study of the K-electron capture phenomenon by Breit and Knipp (Phys. Rev. **54**, 652 (1938)), taken with the adopted mass of Li^7.

C^{11} is obtained from the adopted value for B^{11} and the mass difference $C^{11}-B^{11}$, taken from the article of Livingston and Bethe (Rev. Mod. Phys. **9**, 373 (1937)).

N^{13} is obtained by using the value for C^{13} and the positron energy of N^{13} given by Lyman (Phys. Rev. **55**, 234 (1939)).

C^{14} comes from Bethe (Phys. Rev. **53**, 314 (1938)).

O^{18} was obtained by averaging the values of Mattauch and Aston quoted in Livingston and Bethe. F^{18} is given relative to this value.

F^{20} comes from the results of Burcham and Smith (Proc. Roy. Soc. **A168**, 176 (1938)). The uncertainty was estimated by the writer.

Cl^{36} and S^{36} are derived as follows: Grahame (Phys. Rev. **54**, 972 (1938)) reported that the radioactivity of Cl^{36} was undetectable. Hence it must have very nearly the same mass as A^{36}. Likewise its mass cannot be greater than that of S^{36} by more than $c\cdot1.5$ mMU, since S^{36} is stable. (See A. O. Nier, Phys. Rev. **53**, 282 (1938).) The argument for the isobars at mass 40 is somewhat the same. Henderson (Phys. Rev. **55**, 238 (1938)) puts the upper limit on the beta-rays from K^{40} at 1.3 Mev. K^{40} also goes to A^{40} by K capture. No positrons are observed. From the known mass of A^{40} the other masses are thus obtained.

Sc^{45} is deduced from the results of Pollard (Phys. Rev. **54**, 411 (1938)). When the value given in this reference was put on the $-E_0-\Xi'L$ graph it became clear that something was amiss, because it came out about 9 mMU above its predicted position. The source of the discrepancy was found to be an erroneous quotation of Dempster's value for Ti^{48} in the computation of Sc^{45}. The figure was therefore recomputed and an estimate of the uncertainty made.

Ti^{48} was obtained from the mass-spectrographic measurement of Dempster (Phys. Rev. **53**, 64 (1938)).

V^{51} comes from the transmutation experiment of Davidson and Pollard (Phys. Rev. **54**, 408 (1939)), the uncertainty being estimated by the writer.

Ti^{51} and Cr^{51} come from privately communicated level schemes obtained by H. J. Walke to whom the writer is much obliged for this information.

All the remaining masses were obtained from Livingston and Bethe (Rev. Mod. Phys. **9**, 373 (1937)).

The actual work of this paper was carried through with the earlier value of 1.00813 ± 2 for the mass of H^1. This affects the curve for E_0 slightly (about 0.2 mMU for the heaviest atoms), but has no appreciable effect on the masses.

It has been considered fairly certain that O^{16} completes the $2p$ shell. The present evidence points in that direction, too. One may note in Fig. 1, for example, that the last neutron in O^{17} is bound very loosely. Such an effect is not, however, as evident in the case of Ne^{21}, but the

mass of F^{20} indicates a loss of binding energy of 2.1 mMU for its last neutron, presumably because it is excluded by the Pauli principle from a more tightly bound configuration.

The form of the curve for $-E_0$ suggests the hypothesis that the nuclei from O^{16} to Ne^{20} represent $2s$ configurations.[12] Further information which supports this view is the observed spin ($\frac{1}{2}\hbar$) for F^{19}, and its magnetic moment, which is approximately that of the proton. These are to be compared with the spin ($\frac{3}{2}\hbar$), and the somewhat smaller magnetic moment of Na^{23}.

Atomic Masses

From the curves for $-E_0(A)$ and $L(A)$ it is possible to compute the masses of the various isobaric nuclei at each mass number with considerable accuracy. We use the following formula

$$M=1.00853A+0.0008T_z+E_0+\Xi' L \\ -0.000638A^{-1}(A-T_z-1)T_z+0.0264T_z^2/A,$$

where E_0 and L are taken in mass units. Table IV gives a list of masses thus computed. The uncertainties in the computation become large for very low mass numbers, and for large absolute values of the isotopic spin. Additional uncertainties arise in connection with the corrections at the ends of closed shells, for mass numbers above 40 where the curve for E_0 is not known very well, and for the nuclei of mass $4k+2$ where the effect of the spin dependent forces was demonstrated to influence the binding energy by c·2 mMU. The computed masses, therefore, for which one or more of these effects seems to introduce an uncertainty of more than 1 mMU have been bracketed.

The computed masses are derived directly from the functions $-E_0$ and L, in order that all masses, known and unknown, might be deduced from one systematic scheme. One could, however, reduce the uncertainty in a few of the computations somewhat if the masses were derived relative to accurately known isobars. This is because the function $E_0(A)$, except at the end of a shell, depends on A only; its choice does not affect comparisons between isobars. In the computation of the masses for $A>16$ the value of L has been taken from the simple relationship $L=50/A$ mMU.

Above mass number 40 the data are very scarce and not consistent enough to make computations of masses feasible except at mass numbers 45 and 51 where good agreement is obtained with the observed values.

An effort has been made to collect the best available experimental masses with which to make a comparison. These are also listed in Table IV above the computed mass.

The writer wishes to record his large debt to Professor E. Wigner who was consulted on many occasions for advice in connection with this work

[12] In the single particle models long range forces tend to favor low orbital angular momentum. On the other hand, the sequence of levels for very short range forces is $1s$, $2p$, $3d$, $2s$, $4f$, $3p$, $5g$, $\cdots$. The present information gives the sequence $1s$, $2p$, $2s$, $3d$ for actual many-body nuclei. Although the level filling up after the $3d$ shell is completed at Ca^{40} is probably $4f$, one may not exclude $3p$, $4d$ or $3s$ levels without further information. Because of the nearly symmetrical interaction between nuclear particles one may speak of "nuclear shells" for the light elements. Beyond Ca^{40}, however, the isotopic spin is no longer zero for stable nuclei, and one must speak of "neutron shells" and "proton shells" separately, because they do not close simultaneously. In addition, for very heavy nuclei, many of the striking features of light nuclei which depend on the statistics of the proton and neutron become less evident.

Stark-Effect Broadening of Hydrogen Lines

I. Single Encounters

LYMAN SPITZER, JR.*
Harvard College Observatory, Cambridge, Massachusetts

(Received February 25, 1939)

The profile of a hydrogen emission line is investigated on the assumption that the hydrogen atom is perturbed by passing ions. Provided that the ions do not approach to within ten times the average atomic radius the electric field of the ion may be assumed homogeneous and the matrix elements calculated for the usual Stark effect may be used. For a collision at low velocity or with close approach, the adiabatic approximation is valid, and the profile depends solely on the total phase shift produced by the collision. The Fourier analysis of the wave train may be carried through exactly for certain values of this phase shift. Nonadiabatic transitions are examined for the lowest lines, and an exact solution given for this case also. For higher lines a modified Born approximation is used. The Weisskopf collision broadening formula is shown to hold only in certain cases, and the dispersion curve smearing function for impact broadening is shown to lead in some instances to incorrect results.

THE theory of pressure broadening has led to rigorous solutions only for very slow encounters. In this limiting case the distribution over frequency of the intensity of an emission line, known as the profile of the line, is given simply by the statistical distribution of the perturbing potential produced by the passing ions. The analysis has been carried through by Holtsmark[1] and by Margenau[2] for various potential laws.

While this "statistical broadening," following Margenau's nomenclature, is on a rigorous theoretical basis, the corresponding "impact broadening" for fast particles has been less thoroughly treated. Investigations have usually followed one of two methods. The Weisskopf[3] method on the one hand assumes that two wave trains separated by a phase shift greater than unity[4] are incoherent, and that twice the mean time between such encounters may be taken as the effective damping constant. As will be shown below, this assumption is formally incorrect for frequencies sufficiently far removed from the line center; in addition it neglects the effect of the more distant encounters which, in the case of ionic collisions with hydrogen atoms, may be numerically more important. The second method involves smearing the statistical distribution over some kind of impact distribution. This process is legitimate provided the impact distribution can be correctly chosen, but the dispersion curve sometimes used may lead to quite erroneous results. The approximations made by Lenz[5] restrict the validity of his solution to the core of the line.

For the case of a hydrogen atom perturbed by ions it is possible to give a solution for the line profile more nearly exact than any of the previous impact theories. In what follows three assumptions will be made throughout: (1) The matrix elements of the interaction potential between states of different total quantum number may be neglected, and values for the other matrix elements may be computed on the assumption that the atom is in a homogeneous field of strength Ze/r^2, where r is the distance from the nucleus of the atom to the ion. (2) Each collision, or encounter, may be assumed isolated from all the others. (3) The mass of the colliding particle may be taken infinitely large.

The first of these is the only really serious restriction. This corresponds to the assumption of a linear Stark effect, and neglects all exchange forces, quadratic terms in the Stark effect, forces

* National Reseach Fellow.
[1] J. Holtsmark, Ann. d. Physik **58**, 577 (1919).
[2] H. Margenau, Phys. Rev. **43**, 129 (1933); **44**, 931 (1933); H. Margenau and W. W. Watson, Rev. Mod. Phys. **8**, 22 (1936).
[3] V. Weisskopf, Physik. Zeits. **34**, 1 (1933).
[4] See, however, H. Kallmann and F. London, Zeits. f. physik. Chemie **B2**, 207 (1929), where $\pi/2$ is taken as the critical phase shift.
[5] W. Lenz, Zeits. f. Physik **80**, 423 (1933).

caused by polarization and the like. More seriously, it neglects the inhomogeneity of the Coulomb field produced by the ion. Even if R, the distance of closest approach between the ion and the atom, is ten times $\bar{r}_e$, the average distance of the electron from the nucleus, the perturbing electric force may deviate from its average value by as much as 20 percent. The assumption of a uniform field is scarcely valid, then, for R less than 10 $\bar{r}_e$; since $\bar{r}_e$ equals at most $3n^2a/2$, where a is the radius of the first Bohr orbit, we may take

$$R>8.0\times10^{-8}n^2,$$

where n is the total quantum number.

At such distances the exchange integrals will be very small and the exchange forces negligible. The quadratic Stark effect is also small at such distances. If the electric field is e/R^2, the ratio of the linear shift $\Delta\nu_1$ to the quadratic shift $\Delta\nu_2$ is given by

$$\frac{\Delta\nu_1}{\Delta\nu_2}=\frac{24n(k_2-k_1)R^2}{n^4[17n^2-3(k_1-k_2)^2-9m^2+19]a^2}$$

if k_2-k_1 is replaced by $n/2$ and m and 19 are neglected,

$$\frac{\Delta\nu_1}{\Delta\nu_2}\approx\frac{3R^2}{4a^2n^4}=\frac{27}{16}\left(\frac{R}{\bar{r}_e}\right)^2>150$$

if R is greater than 10 $\bar{r}_e$.

The second assumption simplifies the analysis and is actually not restrictive, as will appear in a second paper to be published shortly. The third assumption neglects both the effect of the perturbation on the motion of the ion relative to the atom and the actual uncertainty of the simultaneous position and motion of the ion. If the ion is an electron, this is, of course, appreciable. The effect of the perturbation is clearly negligible when the Stark shift is only a few angstroms. For the closest encounters, however, for which R equals $8\times10^{-8}n^2$, the maximum interaction energy equals the kinetic energy of a particle at a temperature of 1380° $(n-1)/n^3$; but these close encounters are not frequent and do not contribute appreciably to the observed profiles. As to the uncertainty principle, the results of the analysis depend not on v and R separately but on their product $(l+\frac{1}{2})\hbar/m$, where l is the quantum number for total angular momentum. Since l is a definite constant for each collision the uncertainty principle does not affect the results.

In Section I below the adiabatic wave functions are used in an analysis of the profile of a single component of a hydrogen line when the hydrogen atom is perturbed by a passing ion. Deviations from the adiabatic approximation occur for rapid encounters; these are considered in Section II for the case of the first Lyman lines. In Section III the unperturbed wave functions are used in another determination of the line profile; this corresponds to the Born approximation and, unlike the nonadiabatic analysis in Section II, is applicable to transitions between any two perturbed levels. Integration of the effect of different collisions over all distances and velocities to yield the observable profiles will be the subject of a second paper.

I

A change in potential will be adiabatic and will leave unchanged the quantum state of a system if the change occurs sufficiently slowly. In this case the equation for the line profile may readily be found from the usual Dirac radiation equations as used by Weisskopf and Wigner.[6] We may consider a two-state atom with an upper state A and a lower state B, with energies $E_A(t)$ and $E_B(t)$, respectively. If $H(t)$ is the Hamiltonian operator, $E_A(t)$ and $E_B(t)$ are given by the solutions of the equation

$$H(t)\psi_r(t)=E_r(t)\psi_r(t), \qquad (1)$$

where t is simply a parameter. For the case of a hydrogen atom perturbed by a passing ion, the eigenstates of (1) will be the usual hydrogenic states in parabolic coordinates, rotating so that the z axis is always pointing towards the ion.

Let $a(t)$ denote the probability amplitude of the state in which the electron is in A and there are no photons present. Similarly let $b_\rho(t)$ denote the corresponding amplitude of the state in which the electron is in B and a quantum of frequency ν_ρ has been emitted. If the interaction energy between the photon and the atom is

[6] V. Weisskopf and E. Wigner, Zeits. f. Physik **63**, 54 (1930).

taken as a perturbation, the usual equations for $a(t)$ and $b_\rho(t)$ become

$$ihda/dt=E_A(t)a(t)+\sum_\rho j_{AB^\rho}{}^*b_\rho(t), \quad (2a)$$

$$ihdb_\rho/dt=\{E_B(t)+h\nu_\rho\}b_\rho(t)+j_{AB^\rho}a(t), \quad (2b)$$

where j_{AB^ρ} is the nondiagonal matrix element of the interaction energy between the light quantum and the atom.

A solution for $a(t)$ may be assumed of the form

$$a(t)=e^{-\frac{1}{2}\Gamma t}\exp\left\{\frac{-i}{h}\int_0^t E_A(\tau)d\tau\right\}, \quad (3)$$

which is equivalent to the assumption that the sum in (2a) equals $-\frac{1}{2}ih\Gamma a(t)$. Thus $a(0)$ equals unity and $b_\rho(0)$ must vanish. The Γ used here is twice the usual quantity, but is the only reasonable definition in accordance with the usual concept of a damping constant, since the probability that the electron is in A will be given by $|a(t)|^2$, or $\exp -\Gamma t$, from (3). Integrating (2b) for $b_\rho(t)$, and substituting from (3) for $a(t)$, we find

$$b_\rho(t)=-ij_{AB^\rho}/h\exp\left\{-2\pi i\nu_\rho t-\frac{i}{h}\int_0^t E_B(\tau)d\tau\right\}$$
$$\times\int_0^t e^{-\frac{1}{2}\Gamma T}dT\exp\left\{2\pi i\nu_\rho T+\frac{i}{h}\int_0^T(E_B(\tau)-E_A(\tau))d\tau\right\}. \quad (4)$$

The assumption of (3) may be verified by the use of (4) in computing the sum in (2a). The sum becomes essentially a double Fourier integral of $a(t)$ and hence is proportional to $a(t)$.

The line profile $J'(\nu)$ is simply $|b_{\rho(\nu)}(\infty)|^2$, which gives the distribution of emitted light quanta. It is evident from (4) that one obtains $J'(\nu)$ by squaring the absolute value of the Fourier integral of the damped wave train whose instantaneous frequency is $\{E_A(t)-E_B(t)\}/h$. This is the same result as obtained on the classical theory. If $E_B(t)$ is set equal to zero, $J'(\nu)$ is the absolute value squared of the Fourier transform of $a(t)$.

The result may be put in a more convenient notation. Let the unperturbed energies of A and B be E_{A0} and E_{B0}, respectively, and let the unperturbed frequency be $\nu_{AB}=(E_{A0}-E_{B0})/h$. Then let

$$\Delta_A(t)=\{E_A(t)-E_{A0}\}/h, \quad (5a)$$

$$\Delta_B(t)=\{E_B(t)-E_{B0}\}/h, \quad (5b)$$

$$x=2\pi(\nu-\nu_{AB}). \quad (6)$$

Here x represents the distance in angular frequency units from the unperturbed line center, while Δ_A and Δ_B represent the perturbations to the energy levels, again in angular frequency units. We set $J'(\nu(x))$ equal to $J(x)$ and the square of $|b_\rho(\infty)|$ in (4) becomes in this notation,

$$J(x)=\frac{\Gamma}{2\pi}\left|\int_0^\infty e^{i(x+\frac{1}{2}i\Gamma)T}dT\right.$$
$$\left.\times\exp\left\{-i\int_0^T(\Delta_A(\tau)-\Delta_B(\tau))d\tau\right\}\right|^2, \quad (7)$$

where the integral over all x of $J(x)dx$ has been normalized to unity.

When states A and B are hydrogenic, we have the formula for the linear Stark effect

$$\Delta_A(t)=3hn_A(k_{2A}-k_{1A})F/2me, \quad (8)$$

where n_A is the total quantum number of the Ath state and k_{2A} and k_{1A} are the quantum numbers for the η and ξ coordinates, respectively; F is the electric intensity. If F is produced by an ion of charge Ze at a distance $r(t)$, and if the velocity of the ion is v, its path is a straight line, and its distance of closest approach R is reached at a time t_0, then

$$F=Ze/r^2(t), \quad (9)$$

$$r^2(t)=R^2+v^2(t-t_0)^2, \quad (10)$$

$$\Delta_A(t)-\Delta_B(t)=\frac{q/h}{R^2+v^2(t-t_0)^2}, \quad (11)$$

where

$$q/h=3hgZ/2m=1.73gZ, \quad (12)$$

$$g=n_A(k_{2A}-k_{1A})-n_B(k_{2B}-k_{1B}). \quad (13)$$

Also we find from (11), on integrating,

$$\int_{t_0}^T\{\Delta_A(\tau)-\Delta_B(\tau)\}d\tau=\frac{q}{hRv}\tan^{-1}\frac{v(T-t_0)}{R}. \quad (14)$$

The total phase shift produced in the collision is

thus $\pi q/\hbar Rv$. If we set

$$\delta = q/\hbar Rv, \tag{15}$$

and substitute (14) into (7), we find, after integrating by parts, using (11), and assuming that the integral (14) taken from 0 to t_0 is essentially equal to $\frac{1}{2}\pi\delta$,

$$J(x) = \frac{\Gamma}{2\pi}\frac{1}{x^2+(\frac{1}{2}\Gamma)^2}\left| i - \frac{\delta v}{R} e^{-\frac{1}{2}ix\delta} \right. \tag{16}$$
$$\left. \times \int_0^\infty \frac{e^{i(x+\frac{1}{2}i\Gamma)\tau} d\tau \exp\{-i\delta \tan^{-1} v(T-t_0)/R\}}{1+v^2(T-t_0)^2/R^2} \right|^2.$$

The quantity δ is of fundamental importance in the theory and reappears throughout the analysis. It is equal to $1/\pi$ times the total phase shift produced by the encounter. Since v/R is in general much greater than $\frac{1}{2}\Gamma$, a factor $\exp i(x+\frac{1}{2}i\Gamma)t_0$ may be taken outside the integral in (16). Also, unless x is comparable with Γ, the process of averaging over t_0 will remove the cross product when the square of the absolute value is taken in (16). If, furthermore, T is set equal to $\tau+t_0$ and the lower limit of integration is extended from $-t_0$ to $-\infty$, we have

$$J(x) = (\Gamma/2\pi x^2)\left\{1 + e^{-\Gamma t_0}\delta^2 v^2/R^2 \right.$$
$$\left. \times \left| \int_{-\infty}^{+\infty} \frac{e^{ix\tau} \exp\{-i\delta \tan^{-1} v\tau/R\}}{1+v^2\tau^2/R^2} d\tau \right|^2 \right\}. \tag{17}$$

The $(\frac{1}{2}\Gamma)^2$ in the denominator of (16) has been neglected in (17) since this latter formula is valid only for x greater than $\frac{1}{2}\Gamma$ in any case. Introducing the new variables

$$u = v\tau/R, \quad \xi = xR/v, \tag{18), (19}$$

and integrating (17) over t_0, assuming $\Omega(\delta, \xi)$ collisions per second, we have, finally, for a particular δ and ξ,

$$J(x) = \frac{\Gamma}{2\pi x^2}\left\{1 + \frac{\Omega(\delta,\xi)}{\Gamma}\delta^2 f_\delta^2(\xi)\right\}, \tag{20}$$

$$f_\delta(\xi) = \int_{-\infty}^{+\infty} \frac{e^{i(\xi u - \delta \tan^{-1} u)}}{1+u^2} du. \tag{21}$$

There are several inferences to be drawn from (20) and (21). In the first place $f_\delta(\xi)$ clearly goes to zero with increasing ξ and for sufficiently large

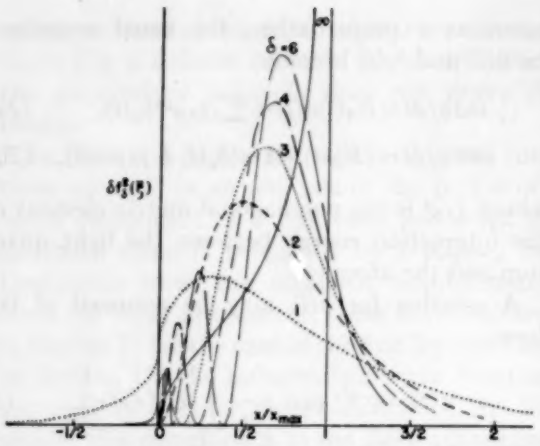


FIG. 1. Line profiles for single encounters. The curves represent $f_\delta^2(\xi)$ for various values of the phase shift $\pi\delta$. Those with odd δ are shown by dotted lines, those with even δ by dashed ones; the solid line represents the statistical distribution; i.e., the limiting case of infinite δ. The abscissae are in units of x/x_{max}, where x is $2\pi(\nu-\nu_{AB})$ and x_{max} is the maximum value of x in the limiting statistical case. The ordinates are proportional to $[x^2+(\frac{1}{2}\Gamma)^2]J(x)$, where $J(x)$ is the line profile; the constant of proportionality is adjusted to give equal areas under all the curves.

x the line profile is quite unaffected by the collisions. This is contrary to statements that the effect of collisions with δ greater than some critical value is simply to split the line into noninterfering parts, each with a much greater damping constant. In the second place $\delta f_\delta(0)$ equals $2\sin(\pi\delta/2)$, as is evident from (21) with the substitution $\tan^{-1}u = \theta$; or, if we average over δ, $\delta^2 f_\delta^2(0)$ equals two. Thus for sufficiently small ξ, Γ is replaced by $2\Omega(\delta, \xi)$, summed over all δ sufficiently large. If $\Omega_{1/\pi}$ is the total number of collisions for which δ is greater than $1/\pi$, or the phase shift is greater than unity, this gives

$$J(x) = \Omega_{1/\pi}/\pi x^2, \tag{22}$$

which is Weisskopf's[7] result.

For integral values of δ, (21) may be integrated exactly. Using the formula

$$\exp(-i\tan^{-1}u) = (1-iu)/(1+u^2)^{\frac{1}{2}}, \tag{23}$$

and expanding $(1-iu)^\delta$, one finds

$$f_\delta(\xi) = \int_{-\infty}^{+\infty} \frac{e^{iu\xi} du}{(1+u^2)^{1+\frac{1}{2}\delta}}$$
$$\times\left(1 - i\delta u + \frac{\delta(\delta-1)(-iu)^2}{2!} \cdots + (-iu)^\delta\right). \tag{24}$$

[7] V. Weisskopf, reference 3. Compare also reference 4.

If the integral of the first term in (24) is called $g_\delta(\xi)$, then this auxiliary function can be integrated[8] and

$$g_\delta(\xi)=\frac{2^{\frac{1}{2}(1-\delta)}\pi^{\frac{1}{2}}}{\Gamma(1+\frac{1}{2}\delta)}\xi^{\frac{1}{2}(1+\delta)}K_{\frac{1}{2}(1+\delta)}(\xi), \quad (25)$$

where $K_\nu(\xi)$ is the usual Bessel function of the second kind for imaginary argument. In terms of $g_\delta(\xi)$, $f_\delta(\xi)$ becomes

$$f_\delta(\xi)=g_\delta(\xi)\mp\delta\frac{d}{d\xi}g_\delta(\xi)+\frac{\delta(\delta-1)}{2!}\frac{d^2}{d\xi^2}g_\delta(\xi)\cdots, \quad (26)$$

where the minus signs on alternate terms go with positive ξ and vice versa.

When δ is an even positive integer 2α, (26) reduces to

$$f_{2\alpha}(\xi)=2\pi(-1)^{\alpha-1}\xi e^{-\xi}\,{}_1F_1(1-\alpha, 2, 2\xi) \quad (27)$$

when ξ is positive, and vanishes when ξ is negative. ${}_1F_1(a, b, x)$ is the confluent hypergeometric series. When δ is equal to one and three, respectively, $f_\delta(\xi)$ becomes

$$f_1(\xi)=2\xi\{K_1(\xi)\pm K_0(\xi)\}, \quad (28a)$$

$$f_3(\xi)=\tfrac{2}{3}\xi\{K_1(\xi)(-1\pm 4\xi)+K_0(\xi)(\mp 3+4\xi)\}, \quad (28b)$$

where again the upper signs go with positive ξ and vice versa. It is evident from (21) that $f_{-\delta}(\xi)$ equals $f_\delta(-\xi)$.

It is readily shown that the integral of $f_\delta^2(\xi)$ over ξ is independent of δ. From the defining formula (21) and from standard theorems on Fourier transforms it follows that

$$\int_{-\infty}^{+\infty} f_\delta^2(\xi)d\xi=2\pi\int_{-\infty}^{+\infty}\frac{du}{(1+u^2)^2}=\pi^2. \quad (29)$$

In Fig. 1 $f_\delta^2(\xi)$ is shown for δ equal to 1, 2, 3, 4 and 6, together with the limiting case of infinite δ. The abscissae represent ξ/δ, a quantity which equals $x/x_{\max}$, where $x_{\max}$ is $q/\hbar R^2$; thus $x/x_{\max}$ is the ratio of the observed frequency shift $\nu-\nu_{AB}$ to the maximum frequency shift in the limit $\delta\to\infty$. To preserve equal areas under each curve, the ordinates are taken to represent $\delta f_\delta^2(\xi)$. Except possibly for $f_1^2(\xi)$, none of these curves deviate markedly from the limiting curve $\delta f_\delta^2(\xi)$ as δ approaches infinity. This limit may be derived directly from (21) or more simply from the usual formula for statistical broadening. In this case $J(x)\cdot dx$ is proportional to the time during which the perturbing potential $\Delta_A-\Delta_B$ lies between x and $x+dx$, in angular frequency units; hence $J(x)$ is proportional to dt/dx, where x equals $q/\hbar r^2(t)$, from (11). Therefore we have, using (10),

$$\frac{dx}{dt}=-\frac{2q}{\hbar r^3}\frac{v^2(t-t_0)}{r}.$$

Since furthermore,

$$v^2(t-t_0)^2=r^2-R^2=R^2(x_{\max}/x-1),$$

we have, taking out an x^{-2}, see (20), and normalizing to π^2 in accordance with (29),

$$\lim_{\delta\to\infty}\delta f_\delta^2(\xi)=2\pi(x_{\max}/x-1)^{-\frac{1}{2}}. \quad (30)$$

It will be noted that the curves for $f_\delta^2(\xi)$ all fall off as $\xi^a\exp -2\xi$. Thus if values of $f_\delta^2(\xi)$ are to be obtained from $f_\infty^2(\xi)$, the limit in which the statistical theory is valid, by use of a smearing function of one sort or another, one must use a function that falls off at least as fast as $\xi^a\exp -2\xi$ if the result is to be asymptotically correct. The dispersion curve sometimes used does not satisfy this requirement, since this falls off only as ξ^{-2}. The "block" and triangle functions used by Margenau are better approximations from this standpoint.

It is also possible to determine $J(x)$ from a consideration of the perturbed wave functions for the ion.[9] If the perturbing potentials are proportional to $1/r^2$, the perturbed wave functions are Bessel functions and the problem can be handled exactly, without neglecting the effect of the potential on the motion of the ion. In general the method leads to the little-investigated hypergeometric functions of the second kind and is thus of little use. When $\delta/2$ is an integer α, however, the formulae simplify. If M and l are the mass and orbital quantum number, respectively, of the incident ion, and the perturbing potential is q/r^2 for the upper state, and zero for

[8] G. N. Watson, *Theory of Bessel Functions* (Cambridge University Press, 1922), Section 6.16, p. 172.

[9] See W. V. Houston, Phys. Rev. **54**, 884 (1938). This corresponds to his method 3 as opposed to method 2 used above.

the lower state, $J(x)$ is given by (20); if q is positive (corresponding to δ positive and the perturbing force repulsive) $f_\delta(\xi)$ is given by

$$f_{2\alpha}(\xi)=2\pi(-1)^{\alpha-1}\left(1-\frac{2\xi}{l+\frac{1}{2}}\right)^{\frac{1}{2}(l+\frac{1}{2})} \times {}_2F_1(1-\alpha, l+\tfrac{3}{2}+\alpha; 2; 2\xi/l+\tfrac{1}{2}), \quad (31)$$

where ${}_2F_1(a, b; c; x)$ is the ordinary hypergeometric series, in this case a terminating polynomial, and where

$$\xi=x\hbar(l+\tfrac{1}{2})/2E_0, \quad (32)$$

and E_0 is the initial or final energy of the ion for an emission or absorption line, respectively; δ is defined in (15). When ξ is negative, $f_{2\alpha}(\xi)$ vanishes. For l large (31) goes over into (27). A similar formula holds when α is negative. If close encounters were sufficiently important it would be possible to investigate by means of these formulas the effect of interaction energies large comparable with the original ionic energy E_0.

II

It is of importance to investigate the limits within which the change of potential is adiabatic and (20) is a valid approximation. Nonadiabatic processes have been investigated by Güttinger[10] and more rigorously by Schwinger[11] for the case of a rotating magnetic field. For a passing ion the analysis is somewhat similar. The perturbations produced by the radiation field will be neglected, since it is evident from (20) that the value of Γ does not affect the shape of the line provided that x is greater than $\frac{1}{2}\Gamma$ but not so great that the collision term $f_\delta^2(\xi)$ in (20) is negligible. Let $a_r(t)$ be the probability amplitude of the rth state, where as before $\psi_r(t)$ satisfies Eq. (1). In the general equation

$$i\hbar\partial\Psi/\partial t=H(t)\Psi, \quad (33)$$

we let

$$\Psi=\Sigma_r a_r(t)\psi_r(t), \quad (34)$$

and find, multiplying through by $\psi_s^*(t)$ and integrating over space,

$$-\frac{da_s}{dt}=i\frac{E_s(t)}{\hbar}a_s(t)+\Sigma_r k_{sr}a_r(t), \quad (35)$$

where

$$k_{sr}(t)=\int\psi_s^*(t)\frac{\partial\psi_r}{\partial t}(t)d\tau; \quad (36)$$

$d\tau$ is an element of volume.

A more convenient expression is available for $k_{sr}(t)$. Differentiation of (1) yields, after rearrangement,

$$\left(\frac{\partial H}{\partial t}-\frac{dE_r}{dt}\right)\psi_r=(E_r(t)-H(t))\frac{\partial\psi_r}{\partial t}. \quad (37)$$

Multiplying by $\psi_s^*(t)$, integrating, and making use of the fact that $H(t)$ is an Hermitian operator and that ψ_s is orthogonal to ψ_r we find, for $s\neq r$,

$$\left(\frac{\partial H}{\partial t}\right)_{sr}=(E_r(t)-E_s(t))k_{sr}. \quad (38)$$

If we neglect states of different total quantum number, k_{ss} vanishes, since ψ_s is not a function of the scalar electric intensity and for any rotation such that the coordinate system does not rotate about its instantaneous z_e axis, where subscripts e refer to electronic coordinates in the perturbed atom, ψ_s will be orthogonal to its rate of change $\partial\psi_s/\partial t$.[12] Eq. (38) thus suffices to determine all relevant k_{sr}.

If the field intensity is again given by (9) and (10), and the ion is assumed to be moving in the $z_e y_e$ plane, where $y_e=(r_e^2-z_e^2)^{\frac{1}{2}}\cos\varphi_e$, that part of H which varies with time will be $-z_e Ze^2/r^2(t)$, and a simple calculation shows that $\partial H/\partial t$ is given by

$$\frac{\partial H}{\partial t}=\frac{Ze^2v}{r^4(t)}\{2z_e v(t-t_0)-y_e R\}. \quad (39)$$

Since the matrix of z_e is diagonal for states of the same total quantum number, only the term in y_e need be considered. Combining the formula for $E_s(t)$,

$$E_s(t)=-\frac{Ze^2z_{ss}}{r^2(t)}, \quad (40)$$

[10] P. Güttinger, Zeits. f. Physik **73**, 169 (1931).
[11] J. Schwinger, Phys. Rev. **51**, 648 (1937).

[12] This may be proved as follows: To preserve the normalization of ψ_s independently of t, k_{ss} must be purely imaginary; i.e., $k_{ss}^*=-k_{ss}$. If we take the plane $\varphi_e=0$ to be the plane in which the ion is moving, this plane will be stationary, and $\partial\psi_s/\partial t$ will have the same symmetry properties with respect to reflection in this plane as ψ_s. Since ψ_s^* is the reflection of ψ_s in this plane, the same relation will hold between $\partial\psi_s/\partial t$ and $\partial\psi_s^*/\partial t$. Hence k_{ss} is unchanged when its complex conjugate is taken, and is purely real. Therefore k_{ss} must vanish.

with (38) and (39) above, we find

$$k_{sr}=\frac{vR}{r^2(t)}\frac{y_{sr}}{z_{rr}-z_{ss}}, \tag{41}$$

where y_{sr} and z_{ss} are the matrix elements of y_s and z_s, respectively. In formula (40) E_{A0} has been set equal to zero.

If now the variable w is introduced, where

$$\{v^2(t-t_0)^2+R^2\}dw=vRdt, \tag{42a}$$

$$w=\tan^{-1}v(t-t_0)/R, \tag{42b}$$

and if $r^2(t)$ is expressed in terms of $E_s(t)$, from (40), then (35) takes the form

$$\frac{da_s}{dw}+i\delta_{ss}a_s+\sum_{r\neq s}\frac{y_{sr}}{z_{rr}-z_{ss}}a_r=0, \tag{43}$$

where δ_{ss} is defined by

$$\delta_{ss}=-Ze^2z_{ss}/\hbar vR. \tag{44}$$

Note that for the Lyman lines and for all lines with unperturbed lower states, this is the same quantity as the δ defined in (15).

The set of simultaneous first-order equations in (43) have constant coefficients and may be solved in any finite case. In practice y_{sr} vanishes except between states whose magnetic quantum number m differs by plus or minus one. For the first excited state of hydrogen the solution is not complicated. If the eigenfunctions are denoted by $\psi_{k_1,\,k_2,\,m}$, let the states ψ_{100}, ψ_{001}, ψ_{00-1} and ψ_{010} be enumerated as ψ_1, ψ_2, ψ_3 and ψ_4, respectively. Then it follows from the eigenfunctions involved[13] that

$$z_{11}=-z_{44};\quad z_{22}=z_{33}=0;$$

$$y_{12}=y_{21}=y_{13}=\cdots=y_{34}=-\tfrac{1}{2}z_{11}. \tag{45}$$

If δ_{ss} is set equal to δ, (43) becomes

$$da_1/dw+i\delta a_1+\tfrac{1}{2}(a_2+a_3)=0, \tag{46a}$$

$$da_2/dw-\tfrac{1}{2}(a_1-a_4)=0, \tag{46b}$$

and similarly for a_3 and a_4. If a solution is assumed of the form $\exp i\sigma w$, the determinant of the coefficients of the a_s in (46) must vanish and σ is given by

$$\sigma=0,\quad \pm(1+\delta^2)^{\frac{1}{2}}. \tag{47}$$

[13] H. Bethe, *Handbuch der Physik*, Vol. 24/1, p. 297.

It would be possible to obtain explicit expressions for the transition probabilities, but there is no need to pursue the analysis further. It is shown above and in the work of Houston[9] and Weisskopf[3] that when the upper state of a radiative transition is perturbed in any way, the line profile is given by the absolute value squared of the Fourier transform of the state function. From the nature of the solution above it is evident that there will be three components of the lowest hydrogen line: one an unshifted line with σ equal to zero; the second arises from the state function

$$a(t)=a(t_0)\exp\{i(1+\delta^2)^{\frac{1}{2}}\tan^{-1}v(t-t_0)/R\}, \tag{48}$$

and the third is the complex conjugate of the second. Comparing these results with the integral in (16) we see that the breakdown of the adiabatic hypothesis affects the form of the line profile in only one way: $f_3(\xi)$ in (20) is replaced by $f_\sigma(\xi)$, where the functional form of $f_\sigma(\xi)$ is still defined by (21) and where σ is taken equal to $(1+\delta^2)^{\frac{1}{2}}$—the trivial value zero in (47) may be neglected. The relative intensities of the shifted and the unshifted components are investigated in Section III below, where the approximation is made that δ is small, but where the analysis takes into account the effect of transitions among the various lower states and eliminates the quantities y_{sr}.

III

When the potential changes rapidly with the time it becomes convenient to assume as a first approximation that the change in the atomic wave functions in the course of the collision may be neglected. This corresponds to the Bohr approximation for fast collisions, except that the emphasis here is on the radiation emitted rather than on the electrons scattered. As in the case of the nonadiabatic transitions discussed above, one cannot use directly the integral in (7) for the line profile, but must start afresh from the Dirac radiation equations.

If only the upper states of a radiative transition are perturbed by collisions, the analysis is not involved.[9] The inclusion of perturbation terms in the equations for the lower states, however, introduces complications. Let there be two states A_1 and A_2 with the same unperturbed

energy E_{A0}; let there be two lower states B_1 and B_2 with unperturbed energy E_{B0}. Let the diagonal matrix elements of the perturbing ionic field be $\hbar\Delta_{A1}(t)$, $\hbar\Delta_{A2}(t)$, $\hbar\Delta_{B1}(t)$, and $\hbar\Delta_{B2}(t)$; let the non-diagonal terms be $\hbar K_A(t)$ and $\hbar K_B(t)$. Let the probability coefficient of a radiative transition from A_1 to B_1 per unit time be Γ. As usual $a_1(t)$ and $a_2(t)$ will represent the probability amplitudes of states A_1 and A_2, while $b_{1\rho}(t)$ and $b_{2\rho}(t)$ will represent the corresponding amplitudes for the states in which the electron is in B_1 and B_2, respectively, and a photon of frequency ν_ρ has been emitted. Then if x is again defined by (6), the equations for the derivatives with time of the probability amplitudes become[6]

$$da_1/dt = -i\Delta_{A1}(t)a_1 - iK_A(t)a_2 - \tfrac{1}{2}\Gamma a_1, \tag{49a}$$

$$db_{1\rho}/dt = -i(x+\Delta_{B1}(t)b_{1\rho} - iK_B(t)b_{2\rho} - ij_{AB^\rho}a_1(t), \tag{49b}$$

$$db_{2\rho}/dt = -i(x+\Delta_{B2}(t))b_{2\rho} - iK_B{}^*(t)b_{1\rho}, \tag{49c}$$

where j_{AB^ρ} is the matrix element for the radiative transition; its value does not concern us. The equation for $a_2(t)$ is not needed. Except for the terms in K_A and K_B these equations (49) are identical with (2), provided that the notational changes of (5) and (6) are made. If the integral over time of $K_B(t)$ is small, these equations may be solved for $b_{1\rho}$ and $b_{2\rho}$ by successive approximations.

First we have for $a_1(t)$

$$a_1(t) = e^{-\frac{1}{2}\Gamma t}h_1(t), \tag{50}$$

$$h_1(t) = \exp\left(-i\int_0^t \Delta_{A1}(\tau)d\tau\right) \times\left\{h_1(0) - i\int_0^t K_A(T)a_2(T) \times\exp\left(i\int_0^T \Delta_{A1}(\tau)d\tau + \tfrac{1}{2}\Gamma T\right)dT\right\}. \tag{51}$$

This satisfies the boundary condition $a_1(0)$ equal to $h_1(0)$. For $b_{1\rho}(t)$ and $b_{2\rho}(t)$ one obtains the integral formulas

$$b_{1\rho}(t) = -ip_1{}^*(t)\int_0^t p_1(T)dT \times\{j_{AB^\rho}e^{-\frac{1}{2}\Gamma T}h_1(T) + K_B(T)b_{2\rho}(T)\}, \tag{52a}$$

$$b_{2\rho}(t) = -ip_2{}^*(t)\int_0^t p_2(T)dT\cdot K_B{}^*(T)b_{1\rho}(T), \tag{52b}$$

where

$$p_1(t) = \exp\left\{ixt + i\int_0^t \Delta_{B1}(\tau)d\tau\right\}, \tag{53}$$

and similarly for $p_2(t)$.

If now we expand $b_{1\rho}t$ in a series of successive approximations,

$$b_{1\rho}(t) = b_{1\rho}{}^{(0)}(t) + b_{1\rho}{}^{(1)}(t) + b_{1\rho}{}^{(2)}(t)\cdots, \tag{54}$$

and similarly for $b_{2\rho}(t)$, we obtain $b_{1\rho}{}^{(0)}(t)$ from (52a), integrating by parts and neglecting $K_B(t)$. This leads to the formula

$$b_{1\rho}{}^{(0)}(t) = \frac{p_1{}^*(t)j_{AB^\rho}}{x+\frac{1}{2}i\Gamma}\left\{h_1(0) - h_1(t)p_1(t)e^{-\frac{1}{2}\Gamma t} + \int_0^t p_1(T)e^{-\frac{1}{2}\Gamma T}\left(\frac{dh_1}{dT} + i\Delta_{B1}h_1(T)\right)dT\right\}, \tag{55}$$

which satisfies the boundary condition $b_{1\rho}(0)$ equal to zero. Substitution of this into (52b) gives $b_{2\rho}{}^{(1)}(t)$. If we neglect all terms that, when squared, contain products or powers of the K's and the Δ's higher than the second, we have

$$b_{2\rho}{}^{(1)}(t) = \frac{-ip_2{}^*(t)j_{AB^\rho}}{x+\frac{1}{2}i\Gamma}\left\{h_1(0)\int_0^t l^*(T)dT - \int_0^t p_2(T)e^{-\frac{1}{2}\Gamma T}K_B{}^*(T)h_1(T)dT\cdots\right\}, \tag{56}$$

where by definition

$$l(t) = K_B(t)p_1(t)p_2{}^*(t). \tag{57}$$

Similarly,

$$b_{1\rho}{}^{(2)}(t) = \frac{-p_1{}^*(t)j_{AB^\rho}}{x+\frac{1}{2}i\Gamma}h_1(0) \times\int_0^t l(\tau)d\tau\int_0^T l^*(T)dT + \cdots. \tag{58}$$

Other terms in $b_{1\rho}{}^{(2)}(t)$ are of effectively higher order since their cross products with $b_{1\rho}{}^{(0)}(t)$ vanish on the average. Since $b_{2\rho}{}^{(0)}$, $b_{1\rho}{}^{(1)}$ and $b_{2\rho}{}^{(2)}$ all vanish, we have to the second order of approximation,

$$|b_{1\rho}(\infty)|^2+|b_{2\rho}(\infty)|^2=|j_{AB}{}^{\rho}|^2/(x^2+(\tfrac{1}{2}\Gamma)^2)$$

$$\left[|h_1(0)|^2\left\{1+\left|\int_0^\infty l(T)dT\right|^2\right.\right.$$

$$\left.-2\Re\int_0^\infty l(\tau)d\tau\int_0^\tau l^*(T)dT\right\}$$

$$+\left|\int_0^\infty p_1(T)e^{-i\Gamma T}dT\left\{\frac{dh_1}{dT}+i\Delta_{B1}h_1\right\}\right|^2$$

$$\left.+\left|\int_0^\infty p_2(T)e^{-i\Gamma T}dTK_B\ h_1\right|^2\right], \quad (59)$$

since, as above, cross terms containing $p_1(T)$ or $p_2(T)$ cancel out when an average is taken over t_0,—the time of closest approach. The symbol $\Re$ denotes the real part of the double integral. If one takes real and imaginary parts of $l(t)$ in the double integral and integrates by parts, it will be seen that this term cancels the one immediately preceding. If we normalize to unity, take exp $-\frac{1}{2}\Gamma T$ outside the integrals, substitute from (50) and (49a) for dh_1/dT, and finally make use of the fact that the change in h_1 and h_2 in the course of the collision may be neglected, taking both these quantities outside the integral and setting the mean square value of each equal to unity, we have for $J(x)$, when $x\gg\frac{1}{2}\Gamma$,

$$J(x)=\Gamma/2\pi x^2\left[1+\exp(-\Gamma t_0)\right.$$

$$\times\left\{\left|\int_0^\infty e^{+ixT}(\Delta_{A1}-\Delta_{B1})dT\right|^2\right.$$

$$\left.\left.+\left|\int_0^\infty e^{ixT}K_AdT\right|^2+\left|\int_0^\infty e^{ixT}K_B{}^*dT\right|^2\right\}\right]. \quad (60)$$

The integrals of Δ_{B1} and Δ_{B2} in the exponents,—see (53) and (57)—have been dropped, as these would first appear in the fourth order. Ordinarily the mean-square value of h_1 or h_2 would be $\frac{1}{2}$, but to make the results comparable with those of the previous section, each state must be assumed to have unit population. In deriving (60) we have made use of the fact that the phases of h_1 and h_2 will be random in general and that any terms involving cross products will vanish.

The functional form of the Δ's and the K's will depend on the choice of axes. If the z_s axis (again in parabolic coordinates) is taken parallel to the direction of motion of the ion, and the y_s axis is in the plane of motion, $\Delta_{A1}(T)-\Delta_{B1}(T)$ will equal $q\sin\theta/\hbar r^2$, where θ is the angle subtended at the atom by the distance from the ion to its point of closest approach. Since the matrix of z_s is diagonal, K_A and K_B will arise from the matrix of y_s alone and hence K_A will equal $\kappa_A\cos\theta/\hbar r^2$, where κ_A is some constant. A similar formula holds for K_B. Since

$$\sin\theta=v(T-t_0)/r, \quad \cos\theta=R/r, \quad (61)$$

we have, using (10) and (15),

$$\int_0^\infty e^{ixT}(\Delta_{A1}-\Delta_{B1})dT=\delta\int_{-\infty}^{+\infty}e^{i\xi u}\frac{udu}{(1+u^2)^{\frac{3}{2}}}, \quad (62)$$

$$\int_0^\infty e^{ixT}K_AdT=\frac{\kappa_A}{\hbar Rv}\int_{-\infty}^{+\infty}e^{i\xi u}\frac{du}{(1+u^2)^{\frac{3}{2}}}. \quad (63)$$

These integrals have already appeared in I and equal $-id g_1(\xi)/d\xi$ and $g_1(\xi)$, respectively (see (25)). Hence (60) gives

$$J(x)=\frac{\Gamma}{2\pi x^2}\left\{1+4\exp(-\Gamma t_0)\xi^2\right.$$

$$\left.\times\left(\delta^2K_0{}^2(\xi)+\frac{\kappa_A{}^2+\kappa_B{}^2}{\hbar^2R^2v^2}K_1{}^2(\xi)\right)\right\}. \quad (64)$$

Suppose, however, that the z_s and y_s axes are interchanged. The roles of the Δ's and the K's will be precisely reversed, and in (64) $K_0{}^2$ and $K_1{}^2$ will become interchanged. When $J(x)$ is summed over all possible components of a particular hydrogen line n,—weighting $J(x)$ for each component by the oscillator strength of that component—the result must clearly be independent of the choice of axes. The weighted sum of all values of δ^2 must therefore equal the corresponding sum over all $\kappa_A{}^2$ and $\kappa_B{}^2$. If (64) is

going to be summed we may legitimately replace $(\kappa_A^2+\kappa_B^2)/\hbar^2R^2v^2$ by δ^2.

If, as before, there are $\Omega(\delta, \xi)$ collisions per second for some particular values of ξ and δ, $J(x)$ is given by (20) where now $f_0'(\xi)$ replaces $f_\delta(\xi)$, and

$$f_0'(\xi)=2\xi(K_1^2(\xi)+K_0^2(\xi))^{\frac{1}{2}}, \qquad (65)$$

or in other words $f_0'^2(\xi)$ is simply the average over plus and minus values of $f_1^2(\xi)$. This is in accord with the conclusion reached in II that the shape of the line profile is given by the adiabatic approximation provided that δ is replaced by σ or $(1+\delta^2)^{\frac{1}{2}}$. The change in $f_\sigma(\xi)$ as δ goes from one to zero is hence very small. The observed profile will of course vary considerably over this range, since as δ goes to zero, ξ corresponds to larger and larger x (see (19)).

Also it is evident from (52b) that the probability of a transition from B_1 to B_2 is approximately given by the square of the time integral of $K_B(t)$. This is simply $1/2\pi$ times the total collision phase shift, or in this case $2\kappa_B/\hbar vR$. This is usually of the same order as δ (see formula (45)). Thus the change in the atomic wave functions becomes appreciable and the Born approximation becomes invalid in the same region as that in which the adiabatic approach becomes approximately valid; i.e., for δ between one-half and two.

APRIL 15, 1939 PHYSICAL REVIEW VOLUME 55

Supplementary Identifications in the Spectra of K, Ca, Sc and Ti

L. W. PHILLIPS
Department of Physics, University of Illinois, Urbana, Illinois

(Received February 23, 1939)

About 65 lines, belonging to the aluminum-like spectra of K and Ca, and the silicon-like spectra of Ca, Sc and Ti, are reported. All known term values for these ions are listed. Ionization potentials for Ca VII and Sc VIII are estimated as approximately 127 volts and 158 volts, respectively.

WHITFORD,[1] who used Ekefors'[2] list of K and Ca wave-lengths, has published the identifications of a number of lines, part of which belong to those spectra of potassium and calcium which are isoelectronic with Al I and Si I. A previous paper from this laboratory[3] reported

[1] A. E. Whitford, Phys. Rev. **46**, 793 (1934).
[2] E. Ekefors, Zeits. f. Physik **71**, 53 (1931).
[3] P. G. Kruger and L. W. Phillips, Phys. Rev. **52**, 97 (1937).

TABLE I. *Classified lines of* K VII, Ca VIII, Ca VII, Sc VIII *and* Ti IX.

INT.	λ(VAC)	ν(CM⁻¹)	CLASSIFICATION	INT.	λ(VAC)	ν(CM⁻¹)	CLASSIFICATION
	K VII				Ca VIII		
4	403.980	247,537	$s^2p\ ^2P_{3/2}-s^2d\ ^2D_{3/2}$	8*	357.983	279,343	$sp^2\ ^4P_{3/2}-spd\ ^4D_{5/2}$
10	403.783	247,658	$\ ^2P_{3/2}-\ ^2D_{5/2}$	8	357.497	279,723	$\ ^4P_{1/2}\ \ ^4D_{3/2}$
0	402.267	248,591	$sp^2\ ^4P_{1/2}-spd\ ^4D_{3/2}$	12	357.347	279,840	$\ ^4P_{1/2}\ \ ^4D_{1/2}$
4	401.758	248,906	$\ ^4P_{3/2}-\ ^4D_{5/2}$	5	355.692	281,142	$\ ^4P_{3/2}\ \ ^4D_{1/2}$
7	401.534	249,045	$\ ^4P_{3/2}-\ ^4D_{7/2}$	8	354.985	281,702	$\ ^4P_{3/2}\ \ ^4D_{3/2}$
1	400.091	249,943	$\ ^4P_{3/2}\ \ ^4D_{5/2}$	9	354.508	282,081	$\ ^4P_{3/2}\ \ ^4D_{5/2}$
3	399.485	250,322	$\ ^4P_{5/2}\ \ ^4D_{3/2}$	4	353.699	282,726	$\ ^4P_{1/2}\ \ ^4D_{3/2}$
2	398.988	250,634	$\ ^4P_{5/2}\ \ ^4D_{5/2}$	2	353.004	283,283	$\ ^4P_{1/2}\ \ ^4D_{3/2}$
8	398.931	250,670	$s^2p\ ^2P_{1/2}-s^2d\ ^2D_{3/2}$	10	184.160	543.005	$s^2p\ ^2P_{3/2}-s^24s\ ^2S_{1/2}$
0	398.252	251,097	$sp^2\ ^4P_{1/2}-spd\ ^4D_{1/2}$	7	182.713	547.305	$\ ^2P_{1/2}\ \ ^2S_{1/2}$
1	397.674	251,462	$\ ^4P_{1/2}\ \ ^4D_{3/2}$	4	179.510	557,073	$sp^2\ ^4P_{1/2}-sp4s\ ^4P_{3/2}$
7	176.106	567,840	$s^2p\ ^2P_{3/2}-s^24d\ ^2D_{5/2}$	2	179.188	558,073	$\ ^4P_{3/2}\ \ ^4P_{1/2}$
5	175.189	570,812	$\ ^2P_{1/2}\ \ ^2D_{3/2}$	6	178.687	559,637	$\ ^4P_{1/2}\ \ ^4P_{1/2}$

TABLE I.—*Continued.*

INT.	λ(VAC)	ν(CM⁻¹)	CLASSIFICATION
		Ca VIII	
8	178.641	559,782	$^4P_{3/2}$ — $^4P_{3/2}$
4	178.247	561,018	$^4P_{1/2}$ — $^4P_{3/2}$
4	177.889	562,147	$^4P_{3/2}$ — $^4P_{5/2}$
3	144.160	693,674	$3s^23p\ ^3P_{3/2}$ — $3s^24d\ ^3D_{3/2}$
8	144.120	693,867	$^3P_{3/2}$ — $^3D_{5/2}$
5	143.270	697,982	$^3P_{1/2}$ — $^3D_{3/2}$
4	115.107	868,760	$^3P_{3/2}$ — $3s^25d\ ^3D_{3/2}$
2	114.566	872,860	$^3P_{1/2}$ — $^3D_{3/2}$
		Ca VII	
1	640.357	156,163	$3s^23p^2\ ^3P_2$ — $3p^3\ ^3D_3$
5	639.153	156,457	$^3P_2$ — $^3D_2$
1	630.772	158,536	$^3P_1$ — $^3D_1$
3	630.533	158,596	$^3P_1$ — $^3D_2$
2	624.391	160,156	$^3P_0$ — $^3D_1$
2	334.904	298,593	$3s^23p^2\ ^3P_2$ — $3s^23pd\ ^3D_1$
5	334.357	299,081	$^3P_2$ — $^3D_3$
10	334.136	299,279	$^3P_2$ — $^3D_2$
3	332.184	301,038	$^3P_1$ — $^3D_1$
8	331.650	301,523	$^3P_1$ — $^3D_2$
5	330.403	302,661	$^3P_0$ — $^3D_1$
6	205.404	486,845	$3s^23p^2\ ^3P_2$ — $3s^23p4s\ ^3P_1$
0	204.756	488,385	$^3P_1$ — $^3P_0$
2	204.377	489,292	$^3P_1$ — $^3P_1$
9	204.001	490,193	$^3P_2$ — $^3P_2$
4	203.700	490,919	$^3P_0$ — $^3P_1$
3	202.989	492,637	$^3P_1$ — $^3P_2$
		Sc VIII	
4	494.301	202,306	$3s^23p^2\ ^3P_2$ — $3p^3\ ^3P_2$
2	493.998**	202,430	$^3P_2$ — $^3P_1$
2	486.523	205,540	$^3P_1$ — $^3P_2$
2	486.242	205,659	$^3P_1$ — $^3P_1$
1	486.192	205,680	$^3P_1$ — $^3P_0$
2—	480.926	207,932	$^3P_0$ — $^3P_1$
12	389.892	256,481	$3s^23p^2\ ^1D_2$ — $3p^3\ ^1P_1$
4	362.308	276,008	$3s^23p^2\ ^3P_2$ — $^1P_1$
2	358.107	279,246	$^3P_1$ — $^1P_1$
		Ti IX	
10	341.754	292,608	$3s^23p^2\ ^3P_2$ — $3p^3\ ^5S_2$
4	336.902	296,822	$^3P_1$ — $^5S_2$
8	289.728	345,151	$3s^23p^2\ ^3P_2$ — $3s^23pd\ ^3P_2$
5	286.234	349,365	$^3P_1$ — $^3P_2$
4	286.128	349,494	$^3P_2$ — $^3P_1$
2	282.720	353,707	$^3P_1$ — $^3P_1$
2	281.463	355,286	$^3P_1$ — $^3P_0$

* Blend with Ca III line.
** May be second O line.

TABLE II. *Term values of* K VII *and* Ca VIII.

DOUBLETS	K VII	Ca VIII
$3s^23p\ ^3P_{1/2}$	0	0
$^3P_{3/2}$	3,129	4,305
$3s^23d\ ^3D_{3/2}$	151,882	171,573
$^3D_{5/2}$	152,049	171,828
$3s^23s\ ^3S_{1/2}$	193,079	216,590
$3s^23p\ ^3P_{1/2}$	206,507	231,012
$^3P_{3/2}$	208,434	233,584
$3s^24d\ ^3D_{3/2}$	250,668	282,362
$^3D_{5/2}$	250,787	282,574
$3s^24s\ ^3S_{1/2}$	439,297	547,308
$3s^24d\ ^3D_{3/2}$	570,812	697,981
$^3D_{5/2}$	570,969	698,172
$3s^25d\ ^3D_{3/2}$		872,860
$^3D_{5/2}$		873,070
QUARTETS		
$3s^23p\ ^4P_{1/2}$	0	0
$^4P_{3/2}$	1,145	1,581
$^4P_{5/2}$	2,871	3,942
$3p^3\ ^4S_{3/2}$	193,479	216,176
$3s3d\ ^4D_{1/2}$	251,092	282,725
$^4D_{3/2}$	251,463	283,283
$^4D_{5/2}$	251,778	283,664
$^4D_{7/2}$	251,916	283,782
$3s^23p^2\ ^4P_{1/2}$	451,314	559,650
$^4P_{3/2}$	452,443	561,017
$^4P_{5/2}$	454,375	563,726
Approx. value of $3s^23p^2\ ^4P_{1/2}$ with resp. to $3s^23p\ ^3P_{1/2}$	114,000	128,000

TABLE III. *Term value of* Ca VII, Sc VIII *and* Ti IX.

	Ca VII	Sc VIII	Ti IX
$3s^23p^2\ ^3P_0$	0	0	0±5
$^3P_1$	1,627	2,273	3,097
$^3P_2$	4,070	5,509	7,311
$^1D_2$	21,870	25,034	
$3s3p^3\ ^3D_1$	160,160		
$^3D_2$	160,228		
$^3D_3$	160,527		
$^3P_2$	185,405	207,914	
$^3P_1$		207,934	
$^3P_0$		207,953	
$^1S_0$	245,232	272,418	299,919
$^1P_1$	252,493	281,515	
$3s^23p3d\ ^3P_2$	286,232	319,570	352,462
$^3P_1$	288,169	322,541	356,805
$^3P_0$	289,011	323,674	358,383
$^3D_1$	302,663		
$^3D_2$	303,151		
$^3D_3$	303,349		
$3s^23p4s\ ^3P_0$	490,012	603,512	
$^3P_1$	490,918	604,605	
$^3P_2$	494,264	609,175	
Est. abs. value of $3s^23p^2\ ^3P_0$ with respect to $3s^23p\ ^2P$ of next stage of ionization	1,030,000	1,280,000	1,560,000

the identification of several scandium lines belonging to the same sequences. This report presents a supplementary list (Table I) of 67 lines which can be assigned to K, Ca, Sc and Ti spectra belonging to these sequences. A complete list of term values for the Al-like ions of K and Ca is given in Table II, and for the Si-like ions of Ca, Sc and Ti in Table III.

The $3s^23p^2\ ^3P_0$ level of Ti IX has not been located. The $(\Delta\nu)^{\frac{1}{2}}$ relation enables one to predict the $^3P_0$—$^3P_1$ separation quite accurately, since it is well known in previous ions of the sequence. The predicted separation is 3097±5 cm⁻¹. The term table for Ti IX is constructed by using 3097 cm⁻¹ as the term value for the $3s^23p^2\ ^3P_1$ level; hence the ground level $3s^23p^2\ ^3P_0$ is listed as 0±5 cm⁻¹.

The Arc Spectrum of Vanadium in the Violet

CHARLOTTE E. MOORE
Princeton University Observatory, Princeton, New Jersey

(Received February 25, 1939)

From new measurements between 2173A and 1848A, six new odd terms and 33 miscellaneous odd levels have been found. Almost all lines of intensity 10 and greater have thus been classified, but it has not been possible to fit them all into multiplets. Five of the new terms and one previously known term may form two triads and probably belong to the $(3d)^3\,4s\cdot 5p$ configuration.

AN extensive analysis of the V I spectrum has been published.[1] At that time, observations in the violet were incomplete. Recently, however, the arc spectrum has been photographed in this region and a total of 263 arc lines measured between 2173A and 1848A. Six new terms and 33 new miscellaneous odd levels have been found. These are listed in Table I, in which the notation is continued from the previous work. A colon denotes doubtful levels. One term, o^4F^0, which has been published, is repeated here with slightly revised values. Six of the terms in Table I may possibly form two triads, n^4D^0, o^4F^0, q^4G^0 and m^4D^0, n^4F^0, p^4G^0. They probably belong to the $(3d)^3\,4s\cdot 5p$ configuration, but the array of terms from this configuration is so complicated that no attempt has been made to assign a limit to either triad. The intensities of the combinations between a^4F and these suggested triads are given in Table II.

Table III contains the newly classified lines

TABLE I. *Terms of* V I.

Term	Level	Diff.	Term	Level
$q^4G_{5/2}^0$	47690.5	133.19	$2a_{11/2}^0$	45353.69*
$q^4G_{7/2}^0$	47823.24	190.94	$4_{3/2,\ 5/2}^0$	46322.39:
$q^4G_{9/2}^0$	48014.18	176.86	$5_{5/2}^0$	46500.64
$q^4G_{11/2}^0$	48191.04		$6_{3/2}^0$	46707.18
			$7_{7/2}^0$	47348.14
$o^4F_{3/2}^0$	47801.6	114.3	$8_{7/2}^0$	47615.56
$o^4F_{5/2}^0$	47915.9	223.5	$9_{3/2,\ 5/2}^0$	47682.68
$o^4F_{7/2}^0$	48139.4	189.4	$10_{5/2}^0$	47809.20
$o^4F_{9/2}^0$	48328.8		$11_{5/2}^0$	47925.49:
			$12_{5/2,\ 7/2}^0$	48001.8:
$n^4D_{1/2}^0$	49189.74	94.03	$13_{7/2}^0$	48023.68
$n^4D_{3/2}^0$	49283.77	156.54	$14_{7/2}^0$	48047.63
$n^4D_{5/2}^0$	49440.31	143.78	$15_{5/2}^0$	48070.91
$n^4D_{7/2}^0$	49584.09		$16_{5/2,\ 7/2}^0$	48201.79
			$17_{7/2}^0$?	48289.8
$r^2F_{5/2}^0$	50404.14	135.13	$18_{5/2}^0$	48881.48
$r^2F_{7/2}^0$	50539.27		$19_{5/2}^0$	48964.99
			$20_{3/2}^0$	49000.82
$p^4G_{5/2}^0$	50452.6:	127.0	$21_{5/2}^0$	49302.61
$p^4G_{7/2}^0$	50579.6	162.8	$22_{7/2,\ 9/2}^0$	49341.90:
$p^4G_{9/2}^0$	50742.4	191.2	$23_{7/2}^0$?	50090.28
$p^4G_{11/2}^0$	50933.58:		$24_{5/2}^0$	50130.6
			$25_{7/2}^0$?	50154.35
$n^4F_{3/2}^0$	50909.7	111.5	$26_{5/2,\ 7/2}^0$	50333.59
$n^4F_{5/2}^0$	51021.2	153.3	$27_{3/2}^0$	50355.89
$n^4F_{7/2}^0$	51174.50	192.1	$28_{5/2,\ 7/2}^0$	50438.35
$n^4F_{9/2}^0$	51366.6		$29_{7/2}^0$	50529.67
			$30_{7/2}^0$	50595.73
$m^4D_{1/2}^0$	50976.5:	91.2	$31_{3/2,\ 5/2}^0$	51194.2
$m^4D_{3/2}^0$	51067.7	144.5	$32_{3/2}^0$	51830.69
$m^4D_{5/2}^0$	51212.2:	185.9	$33_{5/2,\ 7/2}^0$	52008.09
$m^4D_{7/2}^0$	51398.1:		$34_{5/2}^0$	55202.44
			$35_{5/2,\ 5/2}^0$	55877.82:

* Observed $g=1.26$ for this level; calculated from 3556A and 3379A, having observed Zeeman patterns (0) 1.27 and (0) 1.08, respectively.

[1] W. F. Meggers and H. N. Russell, Nat. Bur. Stand. J. Research 17, 125 (1936).

TABLE II. *Intensities in multiplets of* V I.

	$a^4F_{9/2}$	$a^4F_{7/2}$	$a^4F_{5/2}$	$a^4F_{3/2}$
$n^4D_{7/2}^0$	90	50	0	
$n^4D_{5/2}^0$		80	40	1
$n^4D_{3/2}^0$			90	40
$n^4D_{1/2}^0$				60
$o^4F_{9/2}^0$	60*r*	30*r*?*		
$o^4F_{7/2}^0$	25	30	30*r*?*	
$o^4F_{5/2}^0$		25	10*r*?*	20
$o^4F_{3/2}^0$			30	20*r*
$q^4G_{11/2}^0$	40			
$q^4G_{9/2}^0$	15	20*		
$q^4G_{7/2}^0$	0	15	20	
$q^4G_{5/2}^0$		1	15	20*
$m^4D_{7/2}^0$	60	50*	—	
$m^4D_{5/2}^0$		60	50*	—
$m^4D_{3/2}^0$			70	12
$m^4D_{1/2}^0$				60
$n^4F_{9/2}^0$	80*r*?	30		
$n^4F_{7/2}^0$	10	60	30	
$n^4F_{5/2}^0$		15	60	40
$n^4F_{3/2}^0$			15	60
$p^4G_{11/2}^0$	50			
$p^4G_{9/2}^0$	0	50		
$p^4G_{7/2}^0$	—	12	40	
$p^4G_{5/2}^0$		—	—	40

* Blend.

TABLE III. *Classified Lines of* V I.

Ref.	λ(air)Å	I(arc) and Temp. Class	Wave No. vac cm^{-1}	Multiplet Designation
4	3915.36	2 II	25533.26	$b^4D_{3/2}-4^0$
4	3863.40*	1	25876.64	$b^4D_{1/2}-6_{3/2}{}^0$
4	3697.98	1	27034.11	$b^4D_{3/2}-q^4G_{7/2}{}^0$
4	3659.45	2 III	27318.74	$b^2G_{7/2}-19_{9/2}{}^0$
7	3645.626	3 II	27422.34	$a^2F_{7/2}-5_{3/2}{}^0$
4	3636.95	1 III	27487.78	$a^2S_{1/2}-q^4G_{3/2}{}^0$
5	3632.964	1	27517.92	$b^2P_{3/2}-6_{3/2}{}^0$
4	3559.28	1	28087.60	$a^2S_{1/2}-17^0$?
7	3556.241	4 II	28111.58	$a^4G_{11/2}-2a_{11/2}{}^0$
5	3529.474*	[2]	28324.77	$b^2H_{9/2}-7_{7/2}{}^0$
6	3503.181	1 III	28537.35	$a^2F_{7/2}-8_{7/2}{}^0$
4	3476.35	1	28757.61	$b^2G_{7/2}-r^2F_{5/2}{}^0$
5	3460.099*	1 IIIA	28892.66	$b^2G_{7/2}-r^2F_{7/2}{}^0$
6	3454.881*	3 IV	28936.30	$b^2G_{9/2}-r^2F_{7/2}{}^0$
6	3442.006*	2 IV	29044.54	$a^2F_{5/2}-15_{5/2}{}^0$
5	3427.486	1 IV	29167.57	$b^2H_{9/2}-q^4G_{11/2}{}^0$
4	3408.46	1 IIA	29330.41	$b^2G_{9/2}-p^4G_{11/2}{}^0$
5	3401.894*	−1	29386.99	$b^4D_{7/2}-25^0$
4	3385.91	1	29525.69	$b^4D_{1/2}-27_{3/2}{}^0$
4	3383.76	1	29544.46	$b^4D_{3/2}-26^0$
4	3379.35	2 IV	29583.00	$b^4F_{9/2}-2a_{11/2}{}^0$
5	3370.196*	1	29663.37	$b^4D_{5/2}-p^4G_{5/2}{}^0$
4	3345.01	tr IV	29886.70	$a^2F_{7/2}-19_{5/2}{}^0$
5	3242.440	0	30832.11	$a^4G_{5/2}-q^4G_{9/2}{}^0$
5	3230.441*	1	30946.62	$a^4G_{3/2}-12^0$
4	3225.63	1 IIIA	30992.76	$a^4G_{5/2}-14_{7/2}{}^0$
5	3211.323*	0	31130.85	$b^2H_{9/2}-25^0$
4	3180.09	1 IV	31436.56	$b^4P_{3/2}-6_{3/2}{}^0$
6	3147.255	8 IV	31764.56	$a^4G_{7/2}-18_{5/2}{}^0$
8	3143.2	1 IV	31805.6	$a^4S_{3/2}-33^0$
8	3109.42	1 IV	32151.03	$b^2H_{9/2}-n^4F_{7/2}{}^0$
8	3108.56	[1] IV	32159.92	$a^4G_{9/2}-22^0$
8	3099.59	2*h* IV	32252.96	$b^4F_{9/2}-13_{7/2}{}^0$
8	3096.04	2 IV	32289.93	$b^4F_{7/2}-q^4G_{9/2}{}^0$
8	3092.85	1 IV	32323.33	$b^4F_{7/2}-14_{7/2}{}^0$ $a^4G_{7/2}-n^4D_{5/2}{}^0$
6	3038.710	10 IV	32899.16	$a^4H_{7/2}-10_{5/2}{}^0$
4	3027.07	2 IV	33025.64	$b^2P_{1/2}-32_{3/2}{}^0$
8	3009.66*	[1] IV	33216.63	$a^4G_{7/2}-26^0$ $b^4F_{5/2}-18_{5/2}{}^0$
6	3003.288	5 IV	33287.16	$a^4G_{7/2}-r^2F_{5/2}{}^0$
8	3001.05	1*h* IV	33312.02	$b^4F_{5/2}-20_{5/2}{}^0$
8	2994.61	2*h* IV	33383.63	$a^4G_{5/2}-28^0$
8	2991.14*	2 IV	33422.35	$a^4G_{7/2}-r^2F_{7/2}{}^0$
8	2979.21	2 IV	33556.15	$a^2D_{3/2}-15_{5/2}{}^0$ $b^2G_{7/2}-34_{5/2}{}^0$
8	2956.57	1*h* IV	33813.14	$b^4F_{9/2}-n^4D_{7/2}{}^0$
8	2949.91*	2*h* IV	33889.46	$a^2P_{3/2}-q^4G_{5/2}{}^0$
8	2941.11*	1 IIIA	33990.83	$a^2P_{1/2}-o^4F_{3/2}{}^0$
8	2939.26	2*h* V	34012.22	$b^4P_{5/2}-n^4D_{7/2}{}^0$
8	2895.16	4*h* V	34530.27	$a^4H_{7/2}-n^4D_{5/2}{}^0$
5	2849.086	4	35088.69	$b^4D_{5/2}-35^0$
6	2815.994	5*h* V	35501.02	$a^2P_{3/2}-21_{5/2}{}^0$
4	2788.16	2 V	35855.41	$a^2D_{5/2}-r^2F_{5/2}{}^0$
4	2785.52	8 IV	35889.37	$a^2D_{5/2}-r^2F_{5/2}{}^0$
4	2777.70	8*h* IV	35990.41	$a^2D_{5/2}-r^2F_{7/2}{}^0$
1	2742.250	2	36455.64	$a^2G_{7/2}-7_{7/2}{}^0$
1	2711.874	1	36863.96	$a^4P_{3/2}-5_{3/2}{}^0$
1	2710.471	2	36883.04	$a^4P_{5/2}-6_{3/2}{}^0$
1	2708.224	2	36913.64	$a^2G_{9/2}-q^4G_{9/2}{}^0$
1	2696.760*	(6) IV	37070.55	$a^4P_{5/2}-6_{5/2}{}^0$
1	2690.065	2	37162.80	$a^4P_{1/2}-6_{3/2}{}^0$
1	2645.343	5	37791.04	$a^4P_{5/2}-8_{7/2}{}^0$
1	2640.684	(6) III**	37857.71	$a^4P_{5/2}-9^0$
1	2629.094	5	38024.59	$a^4D_{3/2}-5_{5/2}{}^0$
8	2614.90	2 III	38230.98	$a^4D_{3/2}-6_{5/2}{}^0$
1	2613.85	2?	38246.34	$a^4P_{5/2}-15_{5/2}{}^0$
8	2602.7	1 III	38410.17	$a^2G_{7/2}-21_{5/2}{}^0$
1	2597.73	1?	38483.66	$a^2G_{9/2}-n^4D_{7/2}{}^0$
1	2559.81	2*h*?	39053.70	$a^2G_{9/2}-25^0$
1	2550.40	1*h*	39197.78	$a^2G_{7/2}-23^0$
1	2549.834	5	39206.48	$a^4D_{3/2}-9^0$
8	2549.62	3 III Fe?	39209.77	$a^4D_{7/2}-11_{5/2}{}^0$
1	2535.835*	1	39422.90	$a^4D_{5/2}-12^0$
1	2535.441	2	39429.02	$a^2G_{9/2}-29_{7/2}{}^0$
1	2534.825	15*h*	39438.61	$a^2G_{9/2}-r^2F_{7/2}{}^0$
1	2532.280	5	39478.24	$b^4F_{7/2}-34_{5/2}{}^0$
1	2531.778	3	39486.07	$a^4D_{7/2}-16^0$
1	2531.20	4*h*	39495.08	$a^2G_{9/2}-30_{7/2}{}^0$
1	2523.505	5*h*	39615.51	$a^4P_{3/2}-n^4D_{3/2}{}^0$
1	2521.615	(3*h*) III	39645.20	$a^4P_{1/2}-n^4D_{1/2}{}^0$
1	2521.512	6*h*	39646.82	$a^4P_{3/2}-n^4D_{3/2}{}^0$ $a^2G_{7/2}-r^2F_{7/2}{}^0$
1	2515.649	6*h*	39739.22	$a^4P_{1/2}-n^4D_{3/2}{}^0$
1	2514.41	10*h*	39758.79	$a^4P_{5/2}-n^4D_{7/2}{}^0$?
1	2482.711	(15*h*) III	40266.39	$a^2G_{9/2}-n^4F_{9/2}{}^0$ $a^4P_{5/2}-23^0$?
1	2481.11	10*h*	40292.37	$a^4H_{7/2}-34_{5/2}{}^0$
1	2455.41	2*h*	40714.07	$a^4D_{5/2}-n^4D_{1/2}{}^0$
1	2454.79	1	40724.35	$a^4D_{5/2}-21_{5/2}{}^0$ $a^4D_{7/2}-n^4D_{5/2}{}^0$
1	2449.78	1	40807.63	$a^4D_{5/2}-n^4D_{5/2}{}^0$
1	2446.150	3*h*	40868.18	$a^4D_{7/2}-n^4D_{7/2}{}^0$
1	2412.81	1?	41432.84	$a^4P_{1/2}-m^4D_{1/2}{}^0$?
1	2407.517	5	41523.93	$a^4P_{1/2}-m^4D_{3/2}{}^0$?
1	2404.544	5*h*	41575.26	$a^4P_{3/2}-m^4D_{5/2}{}^0$ $a^4D_{5/2}-25^0$
1	2231.412	30	44800.71	$a^4F_{9/2}-2a_{11/2}{}^0$
1	2196.56*	2?	45511.47	$a^4D_{7/2}-q^4G_{7/2}{}^0$
1	2172.75	7	46010.1	
1	2170.74†	60*R*	46052.7	$a^4F_{5/2}-r^4G_{5/2}{}^0$
1	2169.85	8	46071.6	
1	2164.88	*15*	46177.4	$a^4F_{7/2}-5_{5/2}{}^0$
5	2164.54	3	46184.6	$a^4F_{5/2}-4^0$
1	2158.12	*15*	46322.0	$a^4F_{5/2}-4^0$
5	2157.80	5	46328.9	
1	2146.64	*10*	46569.7	$a^4F_{5/2}-6_{5/2}{}^0$ $a^6D_{7/2}-18_{5/2}{}^0$
5	2146.28	6	46577.5	
5	2145.20	3	46601.0	
5	2144.13	5	46624.2	$a^4D_{5/2}-34_{5/2}{}^0$
5	2142.45	0	46660.8	$a^6D_{5/2}-18_{5/2}{}^0$
5	2140.33	1	46707.0	$a^4F_{5/2}-6_{5/2}{}^0$
5	2139.45	0	46726.4	$a^4D_{5/2}-34_{5/2}{}^0$
5	2138.62	10	46744.3	$a^4D_{5/2}-19_{5/2}{}^0$
5	2137.71	2	46764.2	
5	2136.27	1*N*‡‡	46795.7	$a^4F_{9/2}-7_{7/2}{}^0$
5	2135.54	1	46811.7	$a^6D_{5/2}-19_{5/2}{}^0$
5	2135.18	0	46819.6	$a^4F_{7/2}-s^2F_{7/2}{}^0$
5	2133.90	0	46847.7	$a^6D_{5/2}-20_{5/2}{}^0$
5	2132.91	8*N*	46869.5	
5	2132.03	3	46888.8	$a^6D_{1/2}-20_{3/2}{}^0$
5	2127.17	5*d*?	46995.9	$a^4F_{5/2}-s^2F_{5/2}{}^0$?
5	2126.74	5	47005.4	$a^4F_{5/2}-s^2F_{7/2}{}^0$
5	2125.84	20	47025.3	$a^4F_{7/2}-7_{7/2}{}^0$
5	2124.15	12	47062.7	$a^4F_{9/2}-8_{7/2}{}^0$
5	2117.48	20‡‡	47210.9	$a^4F_{5/2}-7_{7/2}{}^0$
5	2114.82	0	47270.3	$a^4F_{9/2}-q^4G_{7/2}{}^0$
5	2113.52	0	47299.4	$a^4D_{5/2}-35^0$
5	2111.38	5*N*	47347.3	
5	2110.51	1‡‡	47366.8	$a^4F_{7/2}-q^4G_{5/2}{}^0$
5	2108.97	8	47401.4	$a^4D_{5/2}-35^0$
5	2106.33	15	47460.8	$a^4F_{9/2}-q^4G_{9/2}{}^0$
5	2104.84	20	47494.4	$a^4F_{9/2}-14_{7/2}{}^0$
5	2104.57	15	47500.5	$a^4F_{7/2}-q^4G_{7/2}{}^0$
5	2104.07	3	47511.8	

* Blend. ** Temperature class = that of blend. ‡‡ May be blend with VII.

TABLE III.—*Continued.*

REF.	λ(AIR)Å	I(ARC) AND TEMP. CLASS	WAVE NO. VAC CM^{-1}	MULTIPLET DESIGNATION
5	2103.04	8	47535.1	$a^6D_{3/2}-f^2D_{5/2}{}^0$?
5	2102.58	15	47545.5	$a^4F_{5/2}-9^0$
5	2102.23	15	47553.4	$a^4F_{5/2}-q^4G_{5/2}{}^0$
5	2101.51	*tr*	47569.7	$a^6D_{3/2}-f^2D_{3/2}{}^0$
5	2100.77†	25	47586.4	$a^4F_{9/2}-o^4F_{7/2}{}^0$
5	2100.51†	25	47592.3	$a^4F_{7/2}-o^4F_{5/2}{}^0$
5	2099.55	3	47614.1	
5	2098.50	40	47637.9	$a^4F_{9/2}-q^4G_{11/2}{}^0$
5	2097.36†	30	47663.8	$a^4F_{5/2}-o^4F_{3/2}{}^0$
5	2097.00	8§§	47672.0	$a^4F_{5/2}-10_{5/2}{}^0$
5	2096.72	15	47678.3	$a^4F_{7/2}-12^0$
5	2096.54	*tr*	47682.4	$a^4F_{3/2}-9^0$
5	2096.37	20§§	47686.3	$a^4F_{5/2}-q^4G_{7/2}{}^0$
5	2096.19	20	47690.4	$a^4F_{3/2}-q^4G_{5/2}{}^0$, $a^4F_{7/2}-q^4G_{9/2}{}^0$
5	2095.77	25	47699.9	$a^4F_{7/2}-13_{7/2}{}^0$
5	2095.29	2	47710.9	
5	2095.11	2	47715.0	
5	2094.71	40	47724.1	$a^4F_{7/2}-14_{7/2}{}^0$
5	2094.15	8	47736.8	$a^6F_{9/2}-17^0$
5	2093.70	1	47747.1	$a^4F_{7/2}-15_{5/2}{}^0$
1	2092.44†	60*r*	47775.8	$a^4F_{9/2}-o^4F_{9/2}{}^0$
1	2092.30†	10*r*	47779.0	$a^6D_{7/2}-23^0$, $a^4F_{5/2}-o^4F_{5/2}{}^0$
5	2091.91	4	47787.9	$a^4F_{5/2}-11_{5/2}{}^0$
1	2091.29†	20*r*	47802.1	$a^4F_{3/2}-o^4F_{3/2}{}^0$
5	2090.96	10	47809.7	$a^4F_{3/2}-10_{5/2}{}^0$
5	2090.68†	30	47816.1	$a^4F_{7/2}-o^4F_{7/2}{}^0$
5	2090.54	5	47819.3	$a^6D_{7/2}-24_{5/2}{}^0$
5	2089.94‡	20	47833.0	
5	2089.13	2	47851.5	
5	2088.56	40	47864.6	$a^4F_{5/2}-12^0$
5	2087.97	1	47878.1	$a^4F_{7/2}-16^0$
5	2087.62	10	47886.1	$a^4F_{3/2}-13_{7/2}{}^0$
5	2087.47	8	47889.6	
5	2086.57	15*n* to *r*	47910.2	$a^6D_{5/2}-24_{5/2}{}^0$, $a^4F_{5/2}-14_{7/2}{}^0$
5	2086.33†	20	47915.7	$a^4F_{3/2}-o^4F_{5/2}{}^0$
5	2085.91	20	47925.4	$a^4F_{3/2}-11_{5/2}{}^0$
5	2085.56	10	47933.4	$a^4F_{5/2}-15_{5/2}{}^0$
5	2084.12	10*n* to *v*	47966.5	$a^4F_{7/2}-17^0$
5	2083.65	*tr*	47977.4	$a^6D_{3/2}-24_{5/2}{}^0$
5	2082.84	12*NN*	47996.0	
5	2082.52†	30*r*?	48003.4	$a^4F_{5/2}-o^4F_{7/2}{}^0$, $a^4F_{7/2}-o^4F_{9/2}{}^0$
5	2081.70	0	48022.3	$a^6D_{7/2}-26^0$
5	2081.35	2*N*	48030.4	
5	2080.86	2*N*	48041.7	
5	2080.70	2*N*	48045.4	
5	2080.46	2*N*	48050.9	
5	2079.87	1*N*	48064.5	$a^4F_{5/2}-16^0$
5	2079.56	15*d*? *r*?	48071.7	
5	2078.12	0*N*	48105.0	$a^6D_{9/2}-29_{7/2}{}^0$
5	2077.75	*tr*	48113.6	$a^6D_{5/2}-26^0$
5	2077.16	15	48127.2	$a^6D_{7/2}-28^0$
5	2077.00	5	48131.0	
5	2076.78	0?	48136.0	$a^6D_{5/2}-27_{3/2}{}^0$
5	2076.41	8*N*	48144.6	
5	2073.89	1	48203.1	$a^6D_{3/2}-27_{3/2}{}^0$
5	2073.58	2*N*	48210.3	
5	2073.23	2	48218.5	$a^6D_{5/2}-28^0$, $a^6D_{7/2}-29_{7/2}{}^0$
5	2072.75	10	48229.6	
5	2072.30	6	48240.1	
5	2072.16	0	48243.4	$a^6D_{1/2}-27_{3/2}{}^0$
5	2068.81	5*N* to *r*§	48321.5	
5	2062.78	3	48462.7	
5	2056.11	2	48619.9	
5	2048.99	*tr*	48788.8	$a^4F_{9/2}-22^0$
5	2044.26	2*NN*§§	48901.7	$a^6D_{7/2}-m^4D_{5/2}{}^0$?
5	2043.13	20	48928.7	
5	2041.74	10*N*	48962.0	
5	2041.24	*tr*	48974.0	$a^6D_{9/2}-m^4D_{7/2}{}^0$?, $a^6D_{5/2}-31^0$
5	2041.00	60	48979.8	$a^4F_{7/2}-21_{5/2}{}^0$
5	2039.81	5*N*	49008.4	
5	2039.39	*tr*	49018.4	$a^4F_{7/2}-22^0$
5	2038.85	90	49031.4	$a^4F_{9/2}-n^4D_{7/2}{}^0$
5	2038.45	0 VII?	49041.0	$a^6D_{3/2}-31^0$
5	2035.30	80	49116.9	$a^4F_{7/2}-n^4D_{5/2}{}^0$
5	2034.06	90	49146.9	$a^4F_{5/2}-n^4D_{3/2}{}^0$
5	2033.28	8	49165.7	$a^4F_{5/2}-21_{5/2}{}^0$
5	2032.27	60	49190.2	$a^4F_{3/2}-n^4D_{1/2}{}^0$
5	2032.10	5*N*	49194.3	
5	2029.36	50	49260.7	$a^4F_{7/2}-n^4D_{7/2}{}^0$
5	2028.42	40	49283.5	$a^4F_{3/2}-n^4D_{3/2}{}^0$
5	2027.62	40	49303.0	$a^4F_{5/2}-n^4D_{5/2}{}^0$
5	2025.04	3	49365.8	$a^4F_{7/2}-f^2D_{5/2}{}^0$
5	2021.98	1	49440.5	$a^4F_{3/2}-n^4D_{5/2}{}^0$
5	2021.72	0	49446.8	$a^4F_{5/2}-n^4D_{7/2}{}^0$
5	2020.30	2	49481.6	
5	2018.04	*tr*	49537.0	$a^4F_{9/2}-23^0$
5	2015.41	4	49601.6	$a^4F_{9/2}-25^0$
5	2015.04	0*N* V II?	49610.7	$a^6D_{5/2}-32_{3/2}{}^0$
5	2013.09	3*N*	49658.8	
5	2012.35	20	49677.0	$a^6D_{3/2}-32_{3/2}{}^0$
5	2011.83	3	49689.9	$a^4F_{3/2}-f^2D_{5/2}{}^0$?
5	2011.54	15	49697.0	$a^6D_{7/2}-33^0$
5	2011.33	4*N*	49702.2	
5	2010.66	2 Fe II?	49718.8	$a^6D_{1/2}-32_{3/2}{}^0$
5	2010.48	20	49723.2	$a^4F_{3/2}-f^2D_{3/2}{}^0$?
5	2009.54	8	49746.5	
5	2008.70	10	49767.3	$a^4F_{7/2}-23^0$
5	2007.86	1	49788.1	$a^6D_{5/2}-33^0$
5	2006.12	9	49831.3	$a^4F_{7/2}-25^0$
5	2004.31	4	49876.3	
5	2003.04	8	49907.9	
	λ(VAC)			
5	2000.56	0	49986.0	$a^4F_{9/2}-r^2F_{7/2}{}^0$
5	1995.43	0	50114.6	$a^4F_{7/2}-28^0$
5	1994.34	20*N*	50141.9	
5	1992.46	0	50189.2	$a^4F_{9/2}-p^4G_{9/2}{}^0$
5	1991.75	2	50207.1	$a^4F_{7/2}-29_{7/2}{}^0$?
5	1991.31	*tr*	50218.2	$a^4F_{5/2}-27_{3/2}{}^0$
5	1989.82	12††	50255.8	$a^4F_{7/2}-p^4G_{7/2}{}^0$
5	1989.17	0	50272.2	$a^4F_{7/2}-30_{7/2}{}^0$
5	1988.97	8	50277.3	
5	1987.25	10*N*§§	50320.8	
5	1984.91	50	50380.1	$a^4F_{9/2}-p^4G_{11/2}{}^0$
5	1984.43	6*N*	50392.3	$a^4F_{5/2}-29_{7/2}{}^0$
5	1983.37	50	50419.2	$a^4F_{7/2}-p^4G_{9/2}{}^0$
5	1982.45	40	50442.6	$a^4F_{5/2}-p^4G_{7/2}{}^0$
5	1982.06	40	50452.6	$a^4F_{3/2}-p^4G_{5/2}{}^0$
5	1981.85	1	50457.9	$a^4F_{3/2}-30_{7/2}{}^0$
5	1975.42	10	50622.2	$a^4F_{9/2}-n^4F_{7/2}{}^0$
5	1973.55	5*N*	50670.1	
5	1972.48	15	50697.6	$a^4F_{7/2}-n^4F_{5/2}{}^0$
5	1969.57	15	50772.5	$a^4F_{5/2}-n^4F_{3/2}{}^0$
5	1967.98	80*r*?	50813.5	$a^4F_{9/2}-n^4F_{9/2}{}^0$
5	1966.76	60	50845.1	$a^4F_{9/2}-m^4D_{7/2}{}^0$
5	1966.52	60	50851.3	$a^4F_{7/2}-n^4F_{7/2}{}^0$
5	1965.26	60	50883.9	$a^4F_{5/2}-n^4F_{5/2}{}^0$
5	1965.07	60	50888.8	$a^4F_{7/2}-m^4D_{5/2}{}^0$
5	1964.27	60	50909.5	$a^4F_{3/2}-n^4F_{3/2}{}^0$
5	1963.47	70	50930.2	$a^4F_{5/2}-m^4D_{3/2}{}^0$
5	1961.69	60	50976.5	$a^4F_{3/2}-m^4D_{1/2}{}^0$

† Designation previously published. §§ May be blend with V II.
‡ Multiplet designation previously assigned to this line very doubtful. § V II present. †† Blend with Cu II.

TABLE III.—*Continued.*

Ref.	λ(vac)	I(arc) and Temp. Class	Wave No. Vac cm^{-1}	Multiplet Designation	Ref.	λ(vac)	I(arc) and Temp. Class	Wave No. vac cm^{-1}	Multiplet Designation
5	1959.97	40	51021.2	$a^4F_{3/2}-n^4F_{5/2}^0$	5	1900.00	6	52631.6	
5	1959.36	30	51037.1	$a^4F_{5/2}-n^4F_{7/2}^0$	5	1898.78	5	52665.4	
5	1959.12	30	51043.3	$a^4F_{7/2}-n^4F_{9/2}^0$	5	1897.90	2	52689.8	
5	1958.60	*tr*	51056.9	$a^4F_{5/2}-31^0$	5	1897.49	5*N*	52701.2	
5	1958.18	12	51067.8	$a^4F_{3/2}-m^4D_{3/2}^0$	5	1894.47	8*N*	52785.2	
5	1957.90	50	51075.1	$a^4F_{7/2}-m^4D_{7/2}^0$, $a^4F_{5/2}-m^4D_{5/2}^0$	5	1890.82	3*N*	52887.1	
					5	1890.50	4*N*	52896.1	
5	1947.37	4*N*	51351.3		5	1872.66	5*N*	53400.0	

and all new unclassified lines of intensity greater than 1. From 3915A to 2196A the stronger lines were included in the previously published paper[1] but were unclassified. In column one the numbers refer to the observers, and the key to the numerals is the same as in the older list. Similarly, the other columns contain, respectively, wave-length, arc intensity, temperature class, wave number and multiplet designation. An asterisk following the wave-length denotes that a designation has been published for the line and it is, therefore, a blend. A dagger indicates that the multiplet designation is taken from the earlier list. In the overlapping region of the two lists a few intensities have been revised. These are in italics when the earlier measurements have been retained. Furnace intensities are in brackets. For lines of known temperature class, King's intensities are given unless the intensity is in parentheses, in which case it is that of the observer in column one.

All but four of the classified lines in the newly-measured region arise from the two lowest terms of the atom, a^4F and a^6D. Almost all of the stronger lines have been classified, but it has not been possible to fit all of them into multiplets. Consequently, the reality of some of the numerous miscellaneous levels may be questionable.

There are enough combinations with higher terms to indicate that there is no systematic error in the ultraviolet measurements. Owing to the lack of standards in the ultraviolet, and to the scarcity of combinations in the visible region based on well measured lines, the present wave-lengths are not sufficiently accurate to provide standards, although they suffice to check the reality of nearly all of the new terms.

This paper is an outcome of work on the analysis of V II by W. F. Meggers and the writer. Observations of the spark spectrum in the ultraviolet were urgently needed. Dr. Meggers kindly lent electrodes of pure vanadium and Dr. A. G. Shenstone secured excellent photographs of both arc and spark spectra at the Palmer Laboratory. The writer greatly appreciates this assistance as well as that of Dr. Henry Norris Russell, who suggested this note and has taken a very keen interest in it.

APRIL 15, 1939 PHYSICAL REVIEW VOLUME 55

Time Changes in Emission from Oxide-Coated Cathodes

J. P. BLEWETT
Research Laboratory, General Electric Company, Schenectady, New York

(Received February 1, 1939)

The decay of electron emission which sets in when emission-limited current is drawn from an oxide cathode has been studied as a function of anode voltage, anode current, cathode temperature, and tube temperature. It is concluded that the decay is due to a transport of barium or oxygen by electrolysis and diffusion. A value is deduced for the heat of the diffusion process. Other time variations in emission have been observed and are discussed.

WHEN an emission-limited current is drawn from an oxide-coated cathode, the emission current decays rapidly at first, then more slowly, finally reaching a more or less constant

value. This effect has been observed and discussed by Davisson and Germer,[1] Becker,[2] Becker and Sears,[3] Kniepkamp and Nebel,[4] and others. By most of these authors, the decrease in emission is attributed to an electrolytic deposition of oxygen at the outer surface of the oxide. This should "poison" the active surface either by decreasing the number of active barium atoms or by effecting a complete change in the character of the surface.

Other possible mechanisms suggest themselves, however. For instance, the decrease in activity might be due to electrolytic removal of barium from the surface. Another possibility is that bombardment by positive ions of residual gases may play a significant part. Even at a pressure of 10^{-6} mm of mercury, a cathode emitting a current of 100 milliamp. per sq. cm to an anode one cm away at a potential of 100 volts may be bombarded by as many as 10^{15} positive ions per sq. cm per second. When we consider that a surface layer of barium probably consists of about 10^{15} atoms per sq. cm, it is evident that positive ion bombardment may be

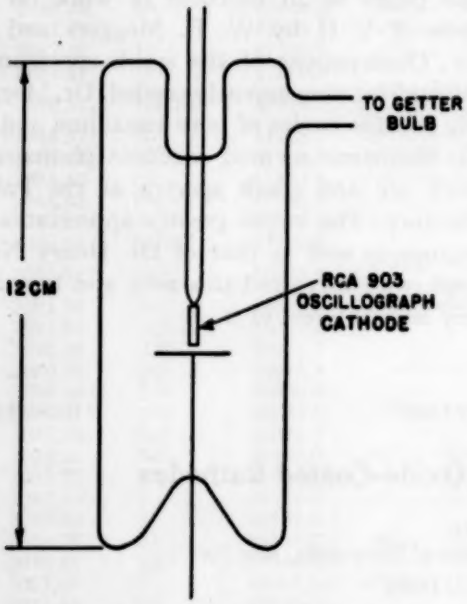


FIG. 1. Experimental tube.

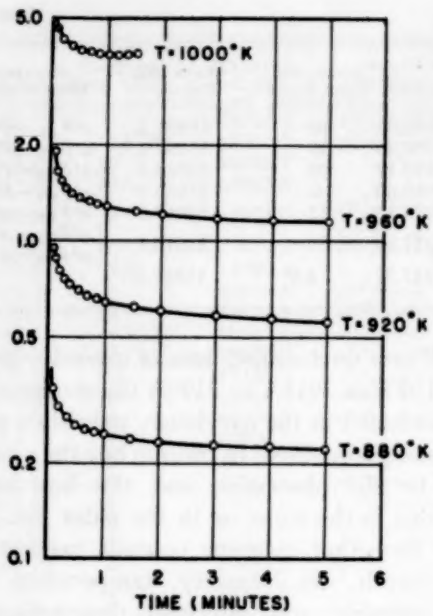


FIG. 2. Decay of emission for various temperatures.

of considerable importance. A third possibility is suggested by Becker's observation[2] that the evolution of oxygen from an oxide cathode depends on whether or not the current is emission-limited, i.e., on the existence of an electric field at the cathode surface. Possibly the decay phenomenon might be caused by a reorientation under the influence of an electric field of surface layers of barium and oxygen ions. Since all of these mechanisms have been proposed as playing an important part in the activation process, it seemed of interest to carry out a more exhaustive investigation of the decay phenomenon.

THE EXPERIMENTAL TUBE

Several commercial tubes with oxide cathodes were tested and all showed the decay effect described above.[5] In the interests of simplicity, a diode was constructed as shown in Fig. 1. The cathode was an RCA type 903 oscillograph cathode. This consists of a nickel can about 3 mm in diameter and 1 cm long closed at one end and indirectly heated by an internal heater. The closed end of the can is coated with a 50–50

[1] C. Davisson and L. H. Germer, Phys. Rev. **24**, 666 (1924).
[2] J. A. Becker, Phys. Rev. **34**, 1323 (1929).
[3] J. A. Becker and R. W. Sears, Phys. Rev. **38**, 2193 (1931).
[4] H. Kneipkamp and C. Nebel, Wiss. Veröff. a.d. Siemens-Konz. **11** (II), 75 (1932).

[5] The effect can easily be observed on the screen of a cathode-ray oscillograph when the current is suddenly increased by raising the potentials of the accelerating electrodes.

percent mixture of BaO and SrO. A flat plate of nickel, separated by 2 mm from the face of the cathode, served as anode. A bulb containing barium getter was connected by a side tube. Following the customary procedure, the tube was baked out, the anode was outgassed by high frequency, the cathode activated, the getter flashed, and the tube sealed off.

A temperature scale was set up for the cathode by pyrometer measurements from 1050 to 1200 degrees K. These readings were corrected for the emissivity of nickel. Lower temperatures were deduced by extrapolation of the log T *vs.* log power input line, corresponding to the equation $T\alpha(\text{power input})^{\frac{1}{4}}$.

The Decay Phenomenon

As noted by Becker,[2] the rate of decay of the emission depends strongly on the cathode temperature. In Fig. 2 we have plotted decay curves for several cathode temperatures. At a temperature of 800 degrees K the decay continues over about an hour. At 1050 degrees K the current has reached a constant value within twenty seconds. At higher temperatures the decay curve can still be traced with an oscillograph, if a square wave of voltage is applied to the anode during a fraction of a cycle of a few milliseconds. The phenomenon is beautifully reproducible so that, although it takes place fifty or a hundred times per second, a clear decay curve unchanged in shape over a period of hours is still visible on the oscillograph screen.

For a fixed temperature the decay effect is independent of anode voltage. The decay curve has been reproduced from 8 to 250 volts on the anode. This would indicate that the effect cannot be attributed to bombardment by positive ions of residual gases. As a further test of the action of residual gases the tube was immersed in liquid air, but no change could be observed in the decay curve.

A marked dependence was observed, however, on the space current. When the current was decreased (by running space charge limited), the range of the phenomenon was decreased so that the final value of the activity was higher than for the higher current. If a high current was passed until the activity reached a constant value and the current was then decreased, it was found that the activity increased to the same value to which it would have decayed had the low current been passed during the whole time. It is evident that the decay process represents an adjustment of the oxide coating to a state of thermodynamic equilibrium depending on the value of the conduction current.

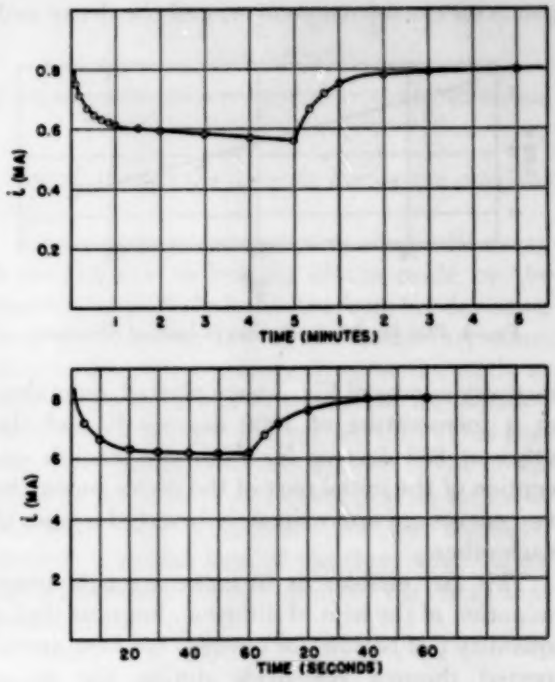


FIG. 3. Decay and recovery of emission at 880°K and 1000°K

The Mechanism of Decay

The dependence of the decay effect on the current through the oxide indicates that the decay phenomenon depends on a volume effect and not a surface effect. The two surface mechanisms which have been proposed, namely, ion bombardment and a reorientation of the surface layers of barium and oxygen under the influence of the field of the anode, are excluded by the independence of anode voltage. Evidently we have to deal with the processes of electrolytic conduction and diffusion within the body of the oxide.

Further light is thrown on the relative parts played by these processes by a study of the rate of recovery of activity when the current is

reduced to zero. To measure the rate of recovery the activity was allowed to decay for a fixed time. The anode was then disconnected, and the cathode was allowed to recover for a measured time. The anode voltage was now turned on, and a decay curve observed. Extrapolation of this curve to zero time gave the activity when the anode voltage was restored. Repetitions of this procedure for several recovery times gave several points on the recovery curve, and the decay and

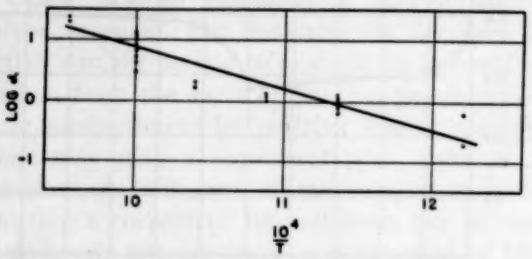


FIG. 4. Plot for determination of heat of diffusion.

recovery curves of Fig. 3 were plotted, one taken at a temperature of 1000 degrees K and the other at 880 degrees K. With the possible exception of the initial part of the decay curve, the two curves are approximately inverted images of each other.

This fact enables us to make a rough determination of the heat of diffusion. Suppose that a quantity q of barium (or oxygen) has been transported through the oxide during the decay process. The concentration gradient responsible for diffusion will be proportional to q, and, during the recovery of activity, since diffusion is the only mechanism, we can write

$$dq/dt = -qD,$$

where D is the diffusion coefficient measured in suitable geometrical units. Hence $q = Ae^{-Dt}$, where A is a constant; but the emission current i is a function of q; therefore $i = i(q) = i(Dt) = i(e^{-E/RT}t)$, where E is the heat of diffusion.

Plotting log i against t, it was found that, by suitable expansion or contraction of the time scales, the decay and recovery curves for different temperatures become superposable within the experimental variation. If α is the factor by which each time scale must be multiplied, it is evident that α is proportional to $e^{-E/RT}$, so that a plot of log α against $1/T$ should give a measure of the heat of diffusion. Approximately the same α seems to apply to decay curves as to recovery curves. This may be taken as indicating either (1) the diffusion is atomic and the heat of diffusion and the temperature coefficient of the electrolytic conductivity are approximately the same; or (2) the diffusion is ionic, in which case the heat of diffusion provides the temperature coefficient for both diffusion and electrolysis.

In Fig. 4 we have plotted log α against $1/T$ for a number of decay and recovery curves. The points scatter rather badly, probably because of variations in state of the oxide from one set of readings to another (for example, the rate of diffusion will be different if a different total quantity of free material is present through the oxide). The observations are represented fairly well by a line whose slope corresponds to a heat of 17 ± 5 kcal. or about 0.7 electron volt.

A point representing the curve given by Becker[6] falls approximately on this line. The point representing the curve given by Becker and Sears[7] is much too high. We should not expect agreement in this case, since this curve applies to a Konel filament where the further process of chemical evolution of barium by reaction with the core must also take part in the decay and reactivation.

Unfortunately it does not seem possible to decide whether barium or oxygen or both are the agents taking part in these processes. It does, however, seem probable that barium makes up at least part of the electrolytic current. From the experiments of Berdennikowa,[8] it is known that barium is liberated electrolytically in amounts proportional to the quantity of charge passed. Becker[2] and others have demonstrated that oxygen also is evolved, but this evolution takes place only when the space current from the oxide is emission-limited so that an electric field exists at the surface of the oxide. This should indicate that the evolution of oxygen is a surface phenomenon rather than a volume transport of oxygen ions by electrolysis.

We are unable also to decide whether the diffusion process is atomic or ionic. Thus it is possible, as mentioned above, that the heat of

[6] Reference 2, Fig. 10.
[7] Reference 3, Fig. 3.
[8] T. P. Berdennikowa, Physik. Zeits. Sowjetunion 2, 77 (1932).

diffusion which has been measured corresponds to the temperature coefficient of the electrolytic conductivity.

Additional Variations in Emission with Time

If the cathode was heated above 1000 degrees K and currents of the order of 30 milliamp. or more per sq. cm were drawn, the emission-limited current was found to vary in a startling manner. After the initial decay the current remained constant for a time, or increased somewhat, then after one or two minutes it began to decay again and finally remained constant at a much lower value. This effect disappeared or was much reduced when the tube was immersed in liquid air. This secondary decay was more pronounced and set in earlier as the current was increased, and it was concluded that it was caused by poisoning by gas, probably oxygen liberated from the anode which was slightly heated by the electron bombardment. Since the anode was initially well outgassed, this gas must have been evaporated from the cathode during the previous runs.

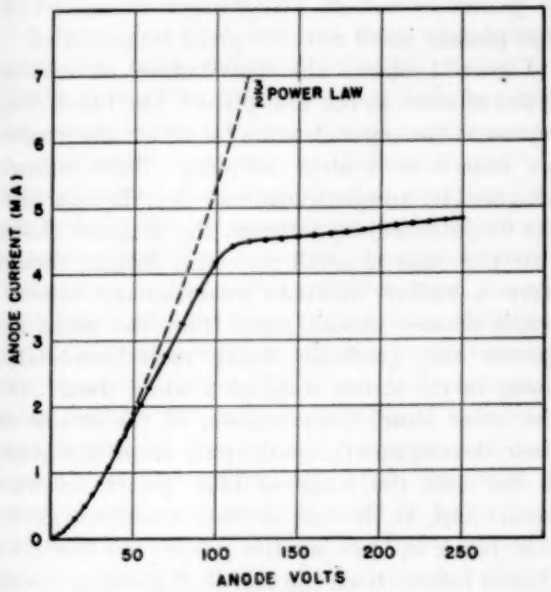


Fig. 5. Emission characteristic at 1000°K.

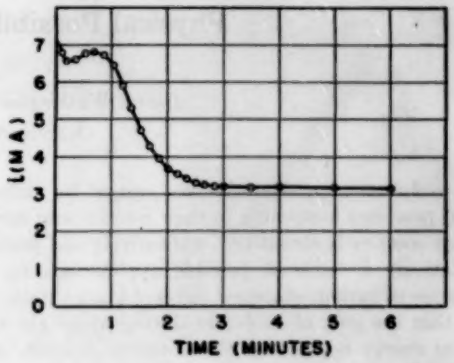


Fig. 6. Combined effects of initial decay, conduction current heating of oxide, and gas from anode.

The increase in current after the initial decay is attributable to heating of the oxide by the emission current. As is evident from the deviation from the $\frac{3}{2}$ power law of the emission characteristic (Fig. 5), a voltage drop in the oxide as high as 10 volts was observed for an emission current of 5 ma (60 ma per sq. cm). This corresponds to an appreciable fraction of the power input to the cathode and would provide sufficient additional heat to account for the increased current. Combinations of the three effects gave rise occasionally to such complex curves as that shown in Fig. 6.

Conclusion

If the initial decay is as fundamental in nature as these observations suggest, it should be a common property of all oxide cathodes. Any change in current drawn from the cathode must result in a gradual change in the cathode activity. Consequently variations in space current from an oxide cathode can be expected to follow voltage variations on other electrodes accurately only if the current from the cathode is space charge limited.

I am indebted to Dr. R. P. Johnson for his collaboration in the performance of several of these experiments, and to Dr. A. W. Hull for many helpful discussions.

Physical Possibilities of Stellar Evolution

G. GAMOW
George Washington University, Washington, D. C.
(Received November 18, 1938)

The evolution of gaseous bodies, caused by different physical processes happening in their interior and serving as energy sources, is considered qualitatively and partially quantitatively in view of possible applications for the explanation of various observed states of known stars. It is shown that the part of evolution during which the main source of energy is given by *thermonuclear reactions* leads to a steadily increasing luminosity and goes over continuously into the *contractive stage* where the energy liberation is purely gravitational. The later stages of contraction and the transition into the state of *degenerate gas model* are discussed, in application to the present state of white dwarfs. Some remarks are made about the possibility of *neutron-core formation* in heavier stars, in application to the explosion phenomena observed in supernovae. An attempt is made to explain the energy production in *red giants* as due to thermonuclear reactions of *light elements* (lithium, beryllium, and boron), and the pulsation phenomena observed for *Cepheid variables* is interpreted as due to instability during the transitions from the giant branch into the main sequence.

I. INTRODUCTION

IT IS well known that a classification of the stars in terms of the mass-radius-luminosity *M-R-L*-relation yields rather simple empirical regularities which evidently are closely related to the internal structure of stars and probably also to their evolutionary history. Most of the known stars belong to the so-called *main sequence;* their luminosities and radii increase rather regularly with their mass. The relation between the radii and luminosities in the main sequence is such that the effective temperature also increases with the mass, so the stars of this group range from hot and luminous *blue giants* down to the cool and faint *red dwarfs*. The stars of comparatively large mass (above $3M(\odot)$)[1] often can be found, however, in a rather different state, in which their luminosities are about the same (or probably somewhat smaller) and the radii much larger than for the stars of the same mass in the main sequence. In such state the stars possess comparatively lower effective temperatures and are called *red giants*. In the *R-L*-diagram these stars seem to form a separate sequence (*giant branch*) branching off near the middle of the main sequence. It should be noticed that some particular stars located in this region of the *R-L*-diagram (Cepheid variables and others of longer periods) possess a property of periodic luminosity changes, usually interpreted as caused by mechanical pulsations of the whole star. In the *R-L*-diagram such pulsating stars represent the upper boundary (in respect to *L*) of the giant branch. Stars of smaller masses, on the other hand, can be found in the so-called *white dwarf* states corresponding to much smaller luminosities and radii than the stars of corresponding masses in the main sequence. Probably related to the white dwarfs are the central stars of planetary nebula (Wolf-Rayet-Stars) which also possess small radii for given luminosities.

Figure 1 shows the distribution of various types of stars in the *R-L*-plane.[2] The black dots represent the separate stars for which the masses are known with some certainty. These masses are given by numbers near each dot.[3] If we accept the evolutionary hypothesis, the diagram shows that the stars of small and large masses should have a rather different evolutionary history. Small masses should start from the main sequence and, gradually losing their luminosity, finish in the dense state of a white dwarf. On the other hand, large masses, in the course of their development, should pass from the stage of red into the stage of blue giants (or vice versa) and, at the end of their evolution, probably break up into smaller pieces; the last conclusion follows from the fact that no dense states

[1] $M(\odot)$ represents the mass of the sun.

[2] We use here the *R-L*-plane, as more suitable for theoretical discussion, instead of the (spectral class)-(absolute magnitude)-plane of the Russell diagram, which refers more closely to direct observational data.

[3] The masses are taken from a critical study by G. P. Kuiper, Astrophys. J. **88**, 472 (1938) which he kindly placed at the disposal of the writer prior to publication.

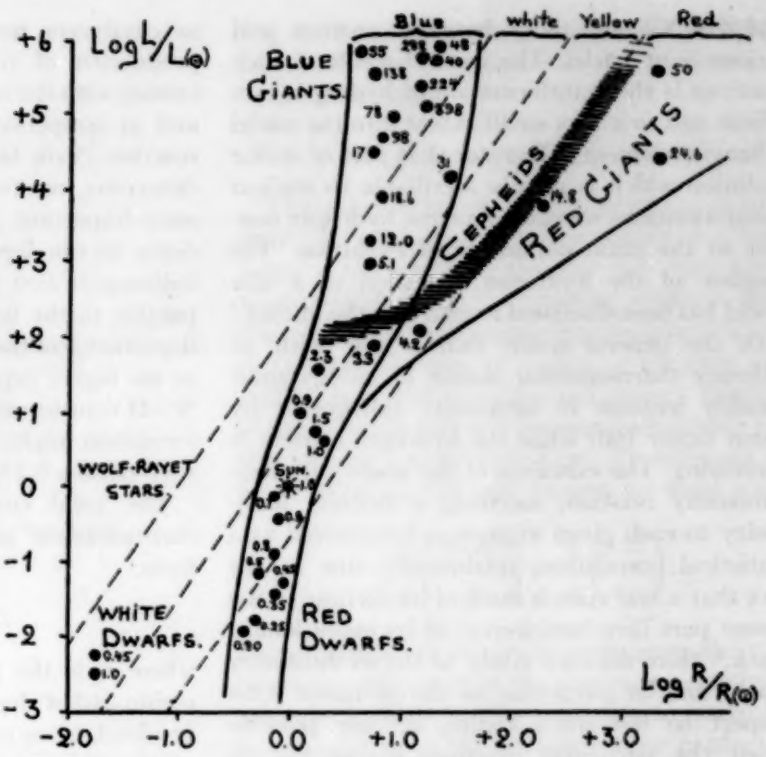


FIG. 1. Showing the relation between masses, radii and luminosities of various stars and the division of stars into the normal stars or the stars of the main sequence, red giants (including Cepheid variables) and white dwarfs (including probably Wolf-Rayet-Stars).

of large masses have ever been observed. Since the hypothesis that different stars were formed originally from different materials seems unlikely, it must be concluded that sufficient time has elapsed since the formation of stars to permit a considerable progress in their evolution.[4] In case of giants, radiating often as much as 1000 erg/g·sec., the preponderant initial hydrogen content would be consumed completely in 10^8 years, i.e., in even shorter time than the age (10^9 years) usually given to the stars. However, for small masses, such as our sun, nuclear energy sources would suffice for about 10^{11} years and, since their formation up to the present time, only one percent of the originally present hydrogen would have been consumed.[5] From this point of view, it seems extremely difficult to understand the existence of white dwarfs which, having a mass comparable to that of the sun and negligible hydrogen-content,[6] require for their formation, through the process of normal evolution, at least 10^{11} years. Perhaps the solution of this difficulty lies in the hypothesis that the white dwarfs known at present do not represent the finite stages of normal evolution of small masses, but are the fragments of much larger stars (giants) which broke to pieces in the finite stage of their rapidly progressing evolution.

In Fig. 1 we observe that the *pulsating stars* (Cepheid variables) are located in a restricted region of *R-L*-diagram, which definitely indicates that the state of pulsation also represents a certain, as yet unknown, stage of the evolution of stars with large masses. The same can be said about the phenomena of *stellar explosions*, such as produce novae and supernovae, for which the regularities observed and the permanent changes in physical state after the explosion also seem to indicate that we are dealing here with discontinuous transitions in the course of stellar evolution.

II. HYDROGEN-EVOLUTION

It is known that the only thermonuclear reactions which can occur at sufficiently large rates at the temperatures of stellar interiors

[4] In fact, equal masses are observed at present in the state of red and blue giants or as the white dwarfs and the stars of the main sequence.
[5] C. v. Weizsäcker, Physik. Zeits. **39**, 633 (1938).
[6] Compare Section IV.

($\geqq 2\times 10^{7}$°C) are those between protons and various light nuclei. The general result of such reactions is the transformation of hydrogen into helium and to a very small extent into the nuclei of heavier elements. Thus for that part of stellar evolution which is mainly ascribable to nuclear transformations we can adopt the hydrogen content as the main parameter of evolution. The problem of the hydrogen evolution of a star model has been discussed recently by the author,[7] with the general result that a star with an ordinary thermonuclear source of energy must steadily increase in luminosity (altogether by about factor 100) while the hydrogen content is decreasing. The existence of the empirical mass-luminosity relation, ascribing a definite luminosity to each given mass, was interpreted as a statistical correlation, intrinsically due to the fact that a star spends most of its lifetime in the lowest part (low luminosity) of its evolutionary track.[8] More detailed study of the evolutionary track, and in particular of the behavior with respect to the star's radius, is only possible when the particular reactions giving rise to the energy production are known. The comparative study of different possible thermonuclear reactions at temperatures around 2×10^{7}°C leads *Bethe*[9] to the conclusion that the energy production of the sun and stars of neighboring mass in the main sequence is exclusively due to a cyclic reaction chain in which the capture of protons by nitrogen and carbon nuclei plays the most important role. The result of such a cyclic reaction chain is the transformation of hydrogen into helium, with negligible formation of heavier nuclei, whereas the amount of N and C acting as catalyzers remains unchanged. The energy production of this reaction chain rapidly decreases with the temperature (as T^{18} at 2×10^{7}°C) and at temperatures below 1.5×10^{7}°C another reaction chain beginning with the formation of deuterons in two-proton collisions[10] becomes more important. The study of astronomical evidence on the dependence of stellar radii on mass indicates in fact that for stars with masses comparable to the mass of the sun the temperature dependence of the energy production corresponds to the higher exponent and thus must be due to N−H transformations, whereas the H−H transformation might be of importance only for stars with masses $0.2M\odot$ and less.[11]

The total energy production caused by a thermonuclear reaction can be written in the form:

$$L=A\times\rho_0^2 r_0^3 T_0^n, \tag{1}$$

where r_0 is the radius of the energy-producing region and A depends on the particular reaction involved, being proportional to the concentration of the element reacting with hydrogen. Combining this equation with the Eddington's fundamental equations of stellar equilibrium and accepting for the opacity dependence on density and temperature the improved expression:

$$\kappa=\kappa_0\rho^{0.5}T^{-2.75} \tag{2}$$

obtained by Bethe on the basis of Strömgren's opacity tables, we come to the following homology transformations for the characteristic data of the star:[12]

[7] G. Gamow, Phys. Rev. **53**, 595 (1938).
[8] G. Gamow, Phys. Rev. **53**, 907 (1938).
[9] H. Bethe, Phys. Rev. **55**, 434 (1939). The author is grateful to Dr. Bethe for being able to see his article before publication.

[10] H. Bethe and Ch. Critchfield, Phys. Rev. **54**, 248 (1938).
[11] G. Gamow, Astrophys. J. **89**, 130 (1939).
[12] These transformations differ from those originally given by the author (Zeits. f. Astrophys. **16**, 113 (1938)) where the Kramers formula was accepted for the opacity changes.

$$\begin{aligned}
&\bar{M}=aM; && \bar{\rho}=a^{-\frac{2n-11.5}{n+1.75}}\,b^{-\frac{3n-20.25}{n+1.75}}\,c^{-\frac{3}{n+1.75}}\,d^{-\frac{3}{n+1.75}}\,\rho,\\
&\bar{\mu}=b\mu; && \bar{P}=a^{-\frac{3n-15.75}{n+1.75}}\,b^{-\frac{4n-27}{n+1.75}}\,c^{-\frac{4}{n+1.75}}\,d^{-\frac{4}{n+1.75}}\,p,\\
&\bar{\kappa}_0=c\kappa; && \bar{T}=a^{+\frac{5}{n+1.75}}\,b^{+\frac{8.5}{n+1.75}}\,c^{-\frac{1}{n+1.75}}\,d^{-\frac{1}{n+1.75}}\,T,\\
&\overline{(Ax)}=d(Ax); && \bar{L}=a^{+\frac{4n+13.25}{n+1.75}}\,b^{-\frac{5.5n+20.25}{n+1.75}}\,c^{-\frac{n+3}{n+1.75}}\,d^{-\frac{1.25}{n+1.75}}\,L,\\
& && \bar{R}=a^{+\frac{n-3.25}{n+1.75}}\,b^{+\frac{n-6.75}{n+1.75}}\,c^{+\frac{1}{n+1.75}}\,d^{+\frac{1}{n+1.75}}\,R.
\end{aligned} \tag{3}$$

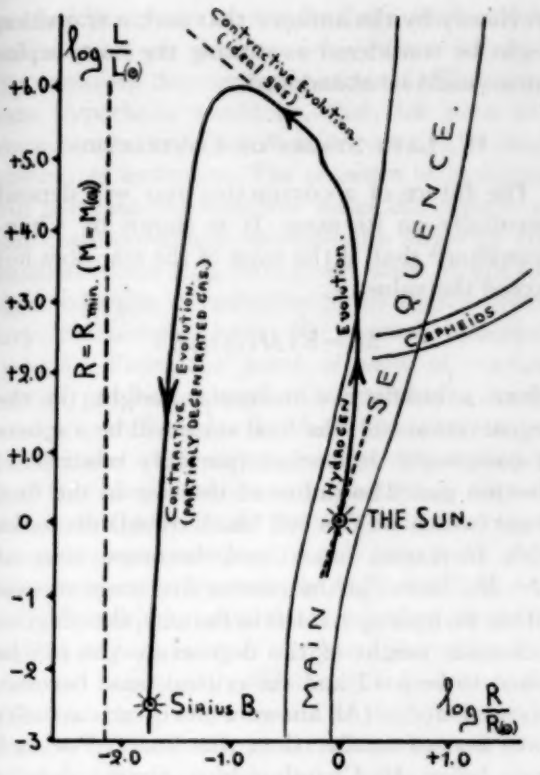


FIG. 2. Evolutionary track of a light star.

From these transformations we can calculate the variations of the physical characteristics of the star with decreasing hydrogen content. For example, for the sun we should start with the observed radius, luminosity, and hydrogen content as estimated for its present state, and use the value 18 for the exponent n. However, with decreasing x the central temperature will gradually increase, which will result in a decrease of n; this should also be taken into account. For very small values of hydrogen content the molecular weight μ and the coefficient of opacity κ_0 become independent of x, and the evolutionary track in the R-L-diagram becomes a straight line. For this part of the evolution we find:

$$L\sim x^{-\frac{1.25}{n+1.75}};\quad R\sim x^{+\frac{1}{n+1.75}}, \tag{4}$$

so that the equation of this line is: $L\sim R^{-5/4}$. The rate of luminosity change with hydrogen content in this stage will also depend essentially on the value of $n(<18)$ at the corresponding temperatures. The evolutionary track based on the N−H reaction is shown schematically in Fig. 2. The broken line passing through the sun represents the dislocation of the track due to variation of mass.

III. Contractive Evolution

The considerations of the previous sections show that, after the hydrogen content has fallen below a certain limit, the star will start a contraction with steadily increasing velocity. When the contraction accompanying the disappearance of the last amounts of hydrogen becomes sufficiently fast, the gravitational energy liberated in such contraction will become comparable with the nuclear energy liberation. To estimate the point at which this will happen we must write:

$$-\frac{d}{dt}\left(\frac{\gamma GM^2}{R}\right)=\epsilon\frac{d}{dt}\left(\frac{xM}{m_H}\right). \tag{5}$$

Here the bracket on the left represents the total gravitational energy of the star (γ is a numerical coefficient which for the point-source model has the value $\frac{3}{5}$), ϵ is the nuclear energy production per proton, and the bracket on the right the total number of protons in the star. From this we get:

$$\frac{d\log R}{d\log x}=Rx\frac{\epsilon}{\gamma GMm_H} \tag{6}$$

or, using (4):

$$\frac{R}{R(\odot)}x=\frac{1}{n+1.75}\frac{\gamma GM(\odot)m_H}{\epsilon}=10^{-5} \tag{7}$$

if we accept $M=M(\odot)$, $\epsilon=2\times10^{-5}$ erg and $n=15$. From this relation and the homology transformation for the sun the gravitational energy liberation becomes predominant when its hydrogen content drops to 0.002 percent and its radius decreases to 0.8 of its present value. However for heavier stars the contractive evolution will set in considerably earlier.

As soon as nuclear energy liberation becomes negligible as compared with gravitational, the energy source ceases to be concentrated in the center, and the energy will be produced throughout the body of the star. The total liberation of gravitational energy inside of a sphere of radius r is evidently given by the change of the total gravitational energy inside this sphere plus the

work of the pressure acting on its surface. Thus we write:[13]

$$L_r=\frac{d}{dt}\left[\frac{\gamma' GM_r^2}{r}\right]-4\pi r^2 p_r\frac{dr}{dt}$$

$$=\left[\frac{\gamma' GM_r^2}{r}+4\pi p_r r^3\right]\left[-\frac{d\log r}{dt}\right] \quad (8)$$

$$-\frac{d\log r}{dt}=S,$$

characterizing the rate of contraction. Choosing M and R as independent variables (now $\mu=$ const.; $\kappa_0=$ const.) we obtain from (8) together with the fundamental equilibrium equations the following homology transformations:

$$\begin{array}{lll}\bar{M}=\alpha M; & \bar{\rho}=\alpha\beta^{-3}\rho; & \bar{L}=\alpha^{5.25}\beta^{-1.25}L,\\ \bar{R}=\beta R; & \bar{P}=\alpha^2\beta^{-4}P; & \bar{S}=\alpha^{3.25}\beta^{-0.25}S.\\ & \bar{T}=\alpha\beta^{-1}T & \end{array} \quad (9)$$

We see that with progressive contraction the luminosity and the rate of contraction both increase. For the relation between luminosity and radius we get the same formula ($L\sim R^{-5/4}$) as for hydrogen evolution with small x. However, for equal radii the contractive model gives a larger luminosity than the corresponding point source model. In fact the calculations of Bierman[14] show that, for the case of a uniform distribution of sources and other conditions equal, the luminosity will be about three times larger than for the case of a point source. Since in the contractive model the source distribution is intermediate between point source and the uniform distribution, we should conclude that the luminosity ratio of the two models for equal radii will be between one and three. The setting in of the contractive equilibrium and the expected corresponding rise of luminosity is indicated schematically in Fig. 2. The detailed changes of the radius and luminosity during the transition is very difficult to calculate (because neither (3) nor (9) are applicable in this stage), but it seems now rather certain that such a transition will take place continuously during a comparatively long period of time. Thus the idea expressed previously by the author[15] that such a transition might be considered as causing the nova explosions should be abandoned.

IV. Late Stages of Contraction

The future of a contracting star will depend essentially on its mass. It is shown by Chandrasekhar[16] that, if the mass of the star does not exceed the value

$$M_0=5.7M(\odot)/\mu^2,$$

where μ is effective molecular weight (in the degenerate state), the final state will be a sphere of completely degenerate (partially relativistic) electron gas. The radius of the star in the final stage (which we can call black dwarf) decreases with increasing mass and becomes zero at $M=M_0$. Since, in the contractive stage of evolution, no hydrogen is left in the star, the effective molecular weight of the degenerate gas can be taken to be $\mu=2$ and the critical mass becomes $M_0=1.4M(\odot)$. (All known white dwarfs actually have masses smaller than this limit.) For each mass below M_0 Chandrasekhar gives a definite value for the minimal radius $R_{\min}$. For $M=M(\odot)$ the value of this radius is $R_{\min}=8\times10^{-3}R(\odot)$. It is shown in Fig. 2 by a vertical dotted line. Thus we see that the contractive evolution (for masses $\leq 1.4M(\odot)$) will not proceed unlimited: even before the intersection of the contractive track for the ideal gas with the line $R=R_{\min}$ deviations should set in, because of the beginning of the formation of a degenerate electron gas in the central regions. The rate of contraction will considerably slow down and the luminosity correspondingly will decrease until it becomes zero and the star turns into a dark inactive body.

Now we can ask how far the known white dwarfs progressed on their way to this finite state. From Fig. 2 we see that the present state of the white dwarf Sirius B (possessing a mass almost equal to that of the sun) is still far away from the limiting line $R=R_{\min}$.[17] Thus we must conclude that this white dwarf has not yet

[13] γ' is numerical coefficient typical for contractive equilibrium, and has the value between 3/2 and 3/5.

[14] L. Bierman, Zeits. f. Astrophys. **3**, 116 (1931).

[15] G. Gamow, Phys. Rev. **54**, 480 (1938).

[16] S. Chandrasekhar, Mon. Not. **91**, 456 (1931); **95**, 207 (1935).

[17] The descending part of the evolutionary track in Fig. 2 is drawn so as to pass through the present position of Sirius B.

reached by a large margin its finite state and its radius must yet decrease by a factor 2.5 until the complete degeneracy is attained. The alternate hypothesis would be, that, for some unknown reason, Sirius B still contains a large amount of hydrogen. The presence of hydrogen will decrease the effective molecular weight μ, and correspondingly increase the value of the minimal radius $R_{\min}$ for a given mass; to get the observed value of radius for Sirius B, it is necessary to accept about 50 percent hydrogen ($\mu=1.5$). From the point of view of nuclear energy production such an hypothesis seems, however, untenable. In fact the analysis of envelopes of white dwarfs leads to the conclusion that the temperature at the bottom of the envelope should be about[18] 2×10^{7}°C and the central temperatures consequently still larger. At such temperatures and extremely high densities, the energy production due to the reaction ${}_1H^1+{}_1H^1 \rightarrow {}_1H^2+\beta^+$ will exceed many million times the observed luminosity of Sirius B. It is very probable that the given estimate of the temperature changes in the envelope are erroneous and that better calculations will give a lower value, but it is doubtful that the calculated temperature can be taken low enough for the luminosity to drop to the observed value. (For this it is necessary to assume the central temperature to be 1.10^{6}°C!) Furthermore, the calculations of Wildhack[19] seem to show that, even accepting the zero-point energy for protons in the central regions of the white dwarfs, one gets too large energy production. Thus it seems that the conclusion that white dwarfs are at present far from the finite stage of contraction is unavoidable.[20]

For stars with masses larger than $1.4M(\odot)$ the process of gravitational contraction is not limited by any maximum density and such stars are apparently destined to unlimited contraction with central density and temperature rising above any given value.

It was first indicated by Sterne[21] that, at very high densities and not-too-high temperatures, the formation of a large number of neutrons must take place because the free electrons are, so to speak, squeezed into the nuclei by the high pressures. In fact we should expect that at very high densities such processes must occur which tend to diminish the number of particles participating in the pressure. If, for example, we consider a mixture of protons and electrons (ionized hydrogen), the process of neutron formation, will reduce the number of particles by a factor of two and we should expect that all electrons will finally be captured by protons if the gain of energy due to the decrease of pressure more than compensates for the energy absorbed in the formation of neutrons. We have seen, however, that very high central densities may be expected only in those stages of stellar evolution which are subsequent to a complete transformation of all available hydrogen into complex nuclei (mostly helium). The transformation of complex nuclei and free electrons into neutrons must there take place according to the formula:

$$ {}_A(\text{nucleus})^Z + Z\bar{e} \rightarrow A\ {}_0n^1, \qquad (10) $$

where A and Z are atomic mass and atomic number of the element in question. Insofar as this corresponds to a transformation of $Z+1$ particles into A particles, the process of neutron formation will lead (except for hydrogen) to the increase of particles participating in pressure and thus will not be in general favored at high densities. However, if the gas is degenerate, the pressure of nuclei can be neglected as compared with pressure due to electrons and the transformation into neutrons will take place. It was

[18] B. Strömgren, Erg. Exakt. Naturwiss. **16**, 465 (1937).
[19] W. Wildhack, Phys. Rev. (in publication).
[20] If we assume that the hydrogen content of white dwarfs is zero we have to explain the strong hydrogen lines in their spectra as due to contamination of hydrogen in their atmospheres. Such discrepancy between internal and external constitution, however, is already known in dealing with several other elements. For example, it has been indicated by Döpel (Zeits. f. Astrophys. **14**, 136 (1937)) that it is necessary to accept central condensation of Li in the sun to be at least 1000 times smaller than in its chromosphere, in order to prevent the liberation of too much energy by the reaction ${}_3Li^7+{}_1H^1 \rightarrow 2{}_2He^4$. Another example is given by the abundance of deuterium, the concentration of which inside of the sun must be entirely negligible according to calculations of Bethe (reference 9). However, if we do not accept a much larger concentration of deuterium in solar atmosphere, it would be difficult to understand its observed abundance on the earth (provided the earth was formed from the sun). The hypothesis that larger concentrations of certain elements in stellar atmospheres is due to the fact that they have been used up in the stellar interior during previous evolution seems the most probable although it contradicts Eddington's conclusion (Mon. Not. **90**, 54 (1929)), that the convection currents should completely mix stellar material once during 10^{7} years.
[21] Th. Sterne, Mon. Not. **93**, 736 (1933); compare also F. Hund, Erg. Exakt. Naturwiss. **15**, 189 (1936).

shown by Chandrasekhar[16] that in the process of the contraction of stars with masses between $5.7M(\odot)/\mu^2$ and $6.6M(\odot)/\mu^2$ the degeneracy of an electron gas will always begin at a certain stage, whereas for heavier stars the central temperatures will rise so rapidly that the electrons will always remain in the state of an ideal gas. Thus we must conclude that the process of neutron formation, accompanied by a rapid decrease of pressure, may happen only for stars of rather sharply defined mass. It was proposed by Baade and Zwicky[22] that a rapid transformation of nuclei into neutrons, causing a collapse of the star with the liberation of an enormous amount of gravitational energy, must be used to explain the vast stellar explosions discovered by them and called "supernovae." The above considerations show that if this explanation is correct the supernovae explosions should take place only for stars with rather sharply defined mass. This seems to be in accord with the observational evidence that all supernovae possess the same luminosity at maximum.

If we ask about the finite state of stars with large masses we find, as indicated above, that no stable finite state does exist and that such stars must undergo continuous unlimited contraction. However it is easy to see that such a process will never continue indefinitely because, since all stars possess an angular momentum, the centrifugal forces will soon become large and will, most probably, cause the breaking of such a massive star into several smaller pieces with the masses below the critical value. These pieces will then continue to exist indefinitely in the form of white dwarfs. Thus we come back to the point of view expressed in the beginning of this article, according to which the existing white dwarfs do not represent a finite stage of the evolution of a single star, but are the fragments resulting from the explosion of heavy stars which have consumed their hydrogen in comparatively short time.

V. The Problem of Red Giants and Cepheid Variables

In the previous sections we have seen that the energy supply of the stars of the main sequence is obtained from the carbon-nitrogen chain reaction and in the later stages of evolution by the gravitational contraction. If we now turn our attention to the stars of the red-giant branch, and in particular to the Cepheid variables, we must come to the conclusion that neither of these two energy sources can be accepted. In fact, because of comparatively low temperatures and densities in the central regions of these stars ($T_0 \leqq 1\times10^{6}$°C; $\rho_0 \leqq 0.001$) the rate of the carbon-nitrogen reaction will be negligibly small, and cannot account for the observed high luminosities. If, on the other hand, we make the hypothesis that the red giants are still in the stage of original contraction (so that no nuclear transformations have yet started), the rate of contraction necessary to explain the observed luminosities comes out too high. In fact, the total gravitational energy U liberated in the contraction from the radius R to $R-\Delta R$ is given by:

$$U \leqq \Delta\left(\frac{GM^2}{R}\right)=\frac{GM^2}{R}\frac{\Delta R}{R}. \qquad (11)$$

For a typical red giant (ζ Aurigae K) we have, for example,

$$M=15M(\odot)=3\times10^{34}\text{g};$$
$$R=200R(\odot)=1.4\times10^{13}\text{ cm}$$

and the total radiation emitted per year $U=7\times10^{44}$ erg. Substituting these values in (11) we get

$$\Delta R/R \leqq 10^{-2} \text{ (per year)}. \qquad (12)$$

Such changes of the radius would cause a secular change of luminosity (0.2 m per decade) which could hardly escape observation.[22a] But perhaps the most striking proof of the constancy of the radius of red giants is given by pulsating stars. According to the observations the period of many such pulsations changes less than by 0.00002 percent per year[23] from which one must conclude, on the basis of general pulsation theory, that the radius of this type of red giant remains constant within approximately the same limit.

[22] F. Baade and F. Zwicky, Proc. Nat. Acad. Sci. 20, 263 (1934).

[22a] *Note added in the proof.*—Professor H. N. Russell kindly informed the author that, according to the ancient observations of Hipparchus, the star ζ Aurigae has not changed by more than one magnitude during the last two thousand years.

[23] P. Bruggencate, Erg. Exakt. Naturwiss. 10, 1 (1931).

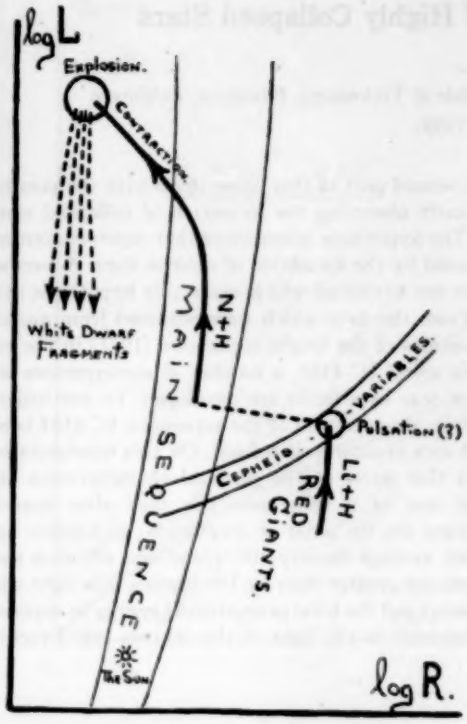


FIG. 3. Evolutionary track of a heavy star.

Thus we are forced to a conclusion that the energy liberation in red giants is due to some nuclear reaction which, however, is going faster and at a lower temperature than the reaction in the stars of the main sequence.

If we exclude the assumption of a resonance phenomenon for some heavier element as too artificial, the only remaining possibility is to accept the assumption that red giants contain in their interior comparatively large amounts of light elements, such as lithium, beryllium and boron, which have already been consumed by the time stars reach the main sequence.[24] Such an assumption becomes still more likely if we remember, as was indicated above, that red giants must be comparatively young stars, formed from interstellar diffuse material long *after* the separation of the great nebula, whereas most of lighter stars (such as the sun) must have been formed *before* the nebular separation.[25]

Since the amount of lithium consumed by nuclear reaction is not renewed (in contrast to the N and C in Bethe's reaction chain), the progressive evolution of a red giant must bring it finally to a stage where, because of a too small lithium-content, the nitrogen reaction becomes of principal importance. The transition of the star into the main sequence of the L-R-diagram (blue-giant region), accompanying this change of the energy producing reaction, must take place continuously but rather rapidly as compared with the total time of evolution; thus we should expect the existence of an "empty-region" in the L-R-diagram between the red and blue giants. A line in this diagram, representing the geometrical place of the points where the carbon-nitrogen reaction becomes predominant for the stars of different masses, should represent an upper limit for the distribution of red giants,[26] and can give rise to an erroneous impression that the giant branch is perpendicular rather than parallel to the main sequence. It should be also indicated here that at the stage of the evolution where the carbon-nitrogen reaction sets in, the curve representing the temperature dependence of the energy liberation must possess a rather sharp bend (because the new reaction varies with a higher exponent of the temperature), and this most probably will give rise to an undamped pulsation lasting until farther evolution carries the star away from this critical point. Long lasting pulsations produced in this way would account for the existence of the Cepheid variables in their peculiar location in the L-R-diagram.

The author is glad to express his thanks to his friends, Dr. H. Bethe and Dr. E. Teller, for valuable discussion.

[24] *Note added in the proof.*—If the energy production in red giants is due to the nuclear reactions of different light elements, these stars must form in the Russell diagram several bands running parallel to the main sequence. The region of pulsating variables is perpendicular to the main sequence, and therefore the relative number of variables with different periods should show maxima at the intersections of this region with the bands corresponding to different light elements. The existence of such maxima is well known: There is a large concentration of stars with periods between 3 and 30 days (ordinary Cepheid variables) and another concentration between 200 and 500 days (long period variables). The estimate of the central temperatures of these two classes of stars leads us to the conclusion that the energy production in ordinary Cepheids is due to the hydrogen reactions with *lithium*, *beryllium*, and *boron* (for different values of periods) whereas the only source of energy in the long period variables is given by the reaction between hydrogen and *deuterium*.

[25] This conclusion concerning the age of stars in great nebula follows from the considerations of the author and E. Teller concerning the formation of the great nebula (Phys. Rev. 55, 654 (1939)).

[26] The calculations of the exact location of this boundary line in the L-R-diagram are at present being carried on by the author and Miss G. White.

On the Theory and Observation of Highly Collapsed Stars

F. ZWICKY
Norman Bridge Laboratory of Physics, California Institute of Technology, Pasadena, California
(Received March 7, 1939)

The investigation presented here falls into two parts. In the first part (Sections A to H) the general relativistic solution given by Schwarzschild of the problem of a homogeneous sphere of constant density is discussed. For every given density a characteristic mass M_l exists, for which the pressure in the center of the sphere becomes infinite, and for which the velocity of light in the same point becomes equal to zero. In the case of collapsed neutron stars M_l is of the order of a large stellar mass. The *effective* mass and the gravitational energy of the sphere are determined as functions of its *proper* mass. Equations are developed which express the velocity and the wave-length of light as functions of the distance from the center of the sphere. The characteristic mass M_l of collapsed neutron stars is expressed in terms of the charge and mass of the electron and the proton and the universal gravitational constant. Some possible relations of the results obtained with recent cosmological theories are pointed out.

The second part of this paper deals with the possibility of actually observing the formation of collapsed neutron stars. The hypothesis is examined that *supernova* outbursts are caused by the formation of *neutron stars*. A number of reasons are advanced which make this hypothesis attractive. From the data which were obtained from numerous observations of the bright supernova (1937) in the extragalactic spiral IC 4182, a number of consequences of the neutron star hypothesis are developed. In particular the *redshift* in the spectrum of the supernova IC 4182 is interpreted as a *gravitational redshift*. On this interpretation it follows that some of the physical characteristics of the central star of a supernova one year after maximum brightness are (in order of magnitude) as follows: radius 100 km, average density 10^{12} g/cm^3 and effective surface temperature greater than 5×10^6 degrees. The light curves, the spectra and the total generation of energy in supernovae are discussed in the light of the neutron star hypothesis.

A. The Neutron Star Hypothesis

THE study of supernovae during the past few years has brought to light many new facts, the correct interpretation of which promises to be of considerable value for the solution of a number of astronomical and physical problems. The remarkable brightness of supernovae and the total energy generated in these stars in the short period of a few weeks made us question some of the current ideas about the generation of energy in stars. As Baade and I have pointed out repeatedly[1,2] it appears doubtful whether any of these ideas can satisfactorily account for the energy liberated in a supernova outburst. It seemed to us in particular that the tremendous *rate* of generation of energy in supernovae would require an explanation along entirely new lines. We suggested[1,2] that such an explanation might be found in the consideration of the transformation of an ordinary star which is composed mainly of electrically charged particles into a *collapsed neutron star* of exceedingly high density (10^{14} g/cm^3) and small stellar radius (10^6 cm). It must be emphasized that we here use the term neutron star simply to designate a highly collapsed star, the average density of which is of the order of the density of matter existing inside of ordinary atomic nuclei. The very properties of space in highly collapsed stars may be radically different from the properties of ordinary space. When we therefore speak of the neutron composition of such a star this does not necessarily mean neutrons in the ordinary sense. It must be rather taken as a short designation for an extended state of matter of nuclear density in which every region whose linear dimensions d are larger than about $\delta=e^2/mc^2$ is essentially electrically neutral, where e and m are the charge and the mass of the electron and c is the velocity of light.

In stellar bodies whose masses are of the same order as the mass of the sun $M(\odot)=2\times10^{33}$ g and whose densities are of the order of 10^{14} g/cm^3 gravitational effects become so large, that most of the other effects usually considered in the theory of stellar constitution become of secondary importance. We therefore disregard in the first approximation the effects of ordinary chemical reactions and even of nuclear reactions and

[1] W. Baade and F. Zwicky, Phys. Rev. **45**, 138 (1934) and **46**, 76 (1934); Proc. Nat. Acad. Sci. **20**, 254 (1934) and **20**, 259 (1934).
[2] F. Zwicky, Sci. Mon. **40**, 461 (1935); Proc. Nat. Acad. Sci., **22**, 457 (1936) and **22**, 557 (1936).

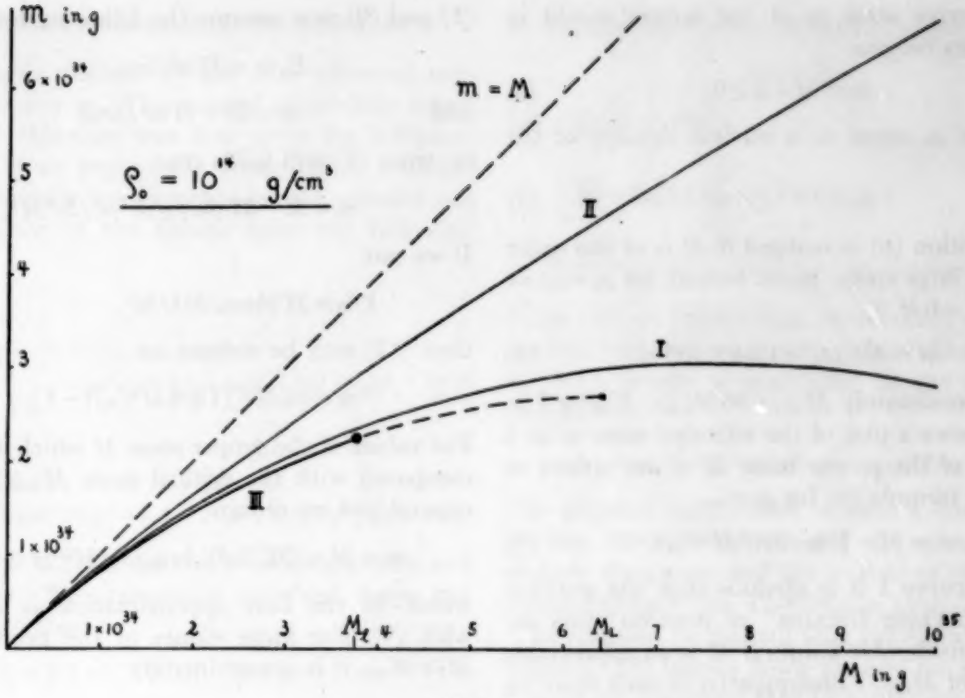


FIG. 1.

limit our discussion essentially to gravitational effects only.

The hypothesis of the formation of neutron stars would run into serious difficulties if one should attempt to retain the classical theory of gravitation. We therefore sketch briefly, why, in the theory of neutron stars it is necessary to introduce general relativistic effects.[3]

B. Difficulties of the Classical Theory of Gravitation

Let us consider, for the sake of illustration, a sphere of radius r_1 and of constant *proper* density ρ_0. The total *proper* mass M of the matter contained in this sphere is, by definition, the mass of this matter when dispersed over a very large volume. The gravitational potential energy E_g of our sphere is

$$E_g = -3\Gamma M^2/5r_1 \tag{1}$$

with

$$M = 4\pi\rho_0 r_1^3/3 \tag{2}$$

and $\Gamma = 6.66\times10^{-8}$ g^{-1} cm^3 sec.$^{-2}$. We rewrite (1) in terms of ρ_0 and M and we obtain

$$E_g = -\tfrac{3}{5}(4\pi/3)^{\frac{1}{3}}\Gamma M^{5/3}\rho_0^{\frac{1}{3}} = -6.44\times10^{-8}M^{5/3}\rho_0^{\frac{1}{3}}. \tag{3}$$

E_g is measured relative to the usual zero configuration in which the mass M is assumed to be dispersed over a very large volume ($\rho=0$). According to the special theory of relativity the mass equivalent of E_g is

$$\Delta = |E_g|/c^2 = 7.2\times10^{-29}M^{5/3}\rho_0^{\frac{1}{3}}. \tag{4}$$

If we indefinitely increase M at constant ρ_0, or ρ_0 at constant M we arrive according to (4) at configurations for which

$$\Delta = M \quad \text{or even} \quad \Delta > M. \tag{5}$$

[3] An outline of the following considerations was first presented in December, 1933, to the Physical Society meeting at Stanford (reference 1). Some details were subsequently published in various papers on supernovae (references 1 and 2). At that time very little was known definitely about supernovae and it seemed premature to discuss in any detail the formation of neutron stars as a possible cause for supernovae. In the meantime, through the discovery with the 18-inch Schmidt telescope on Palomar mountain of eight supernovae, the existence of supernovae as a special class of temporary stars may be regarded as established beyond reasonable doubt (F. Zwicky, Publ. Astr. Soc. Pac. **49**, 204 (1937); **50**, 215 (1938) and **51**, 36 (1939)). Also, the spectral studies of the bright supernovae in IC 4182 (August, 1937) and NGC 1003 (September, 1937) have furnished data (R. Minkowski, Astrophys. J., March (1939); F. Zwicky, Astrophys. J. **88**, 522 (1938)) which fully justify a more detailed examination of the neutron star hypothesis at the present time.

The effective mass m of our sphere would in these cases become

$$m = M - \Delta \lesssim 0. \tag{6}$$

If we set ρ_0 equal to a nuclear density of the order

$$\rho_n = 10^{14}\ \text{g/cm}^3 \tag{7}$$

the condition (6) is realized if M is of the order of a very large stellar mass. Indeed, for $\rho_0 = \rho_n$ we obtain $m = 0$ if

$$M = M_{\text{critical}} = 1.64 \times 10^{33}\ \text{g}. \tag{8}$$

It is approximately $M_{\text{crit}} = 80 M(\odot)$. Curve I in Fig. 1 shows a plot of the effective mass m as a function of the proper mass M of our sphere as given by formula (9) for $\rho_0 = \rho_n$.

$$m = M - \tfrac{3}{5}(4\pi/3)^{\frac{1}{3}} \Gamma M^{5/3} \rho_0^{\frac{1}{3}} c^{-2}. \tag{9}$$

From curve I it is obvious that the gravitational "packing fraction" of neutron stars assumes considerable values if M is an appreciable fraction of M_{crit}. Consequently, in such stars we may expect to encounter conditions with which we are not at all familiar in the theory of the idealized models of ordinary stars which are at present customary.

The question naturally arises whether actual configurations of dense matter in bulk, such as neutron stars, can possibly behave as paradoxially as the relation (9) suggests. Although negative masses have been considered in the theory of elementary corpuscles, we have no observational knowledge of matter in bulk, such as stars, of negative effective mass. We must therefore look for a solution of our problem on the basis of further theoretical considerations. In the first place we may proceed a little farther on the basis of the classical theory of gravitation coupled with the result of the special theory of relativity that the mass equivalent of any type of energy E is $\Delta = E/c^2$. This latter fact suggests that the relation (1) for the gravitational energy of our sphere ceases to be strictly correct as soon as the effective mass m differs from the proper mass M. Although, on the special theory of relativity, we do not know how the mass equivalent Δ of the gravitational energy is to be properly distributed over the sphere, we may, for the sake of illustration smear it uniformly over this sphere, with the result that the relations (1) and (9) now assume the following form

$$E_g = -3\Gamma m^2/5r_1 \tag{10}$$

and

$$m = M - 3\Gamma m^2/5r_1 c^2 \tag{11}$$

or, since (2) still holds true,

$$m = M - 3\Gamma(4\pi\rho_0/3)^{\frac{1}{3}} m^2/5c^2 M^{\frac{1}{3}}. \tag{12}$$

If we put

$$1/\alpha = 3\Gamma(4\pi\rho_0/3)^{\frac{1}{3}}/5c^2, \tag{13}$$

then (12) may be written as

$$m = \tfrac{1}{2}\alpha M^{\frac{1}{3}}[(1 + 4M^{\frac{1}{3}}/\alpha)^{\frac{1}{2}} - 1]. \tag{14}$$

For values of the proper mass M which are small compared with the critical mass M_{crit} we may expand and we obtain

$$m = M - (3\Gamma/5c^2)(4\pi\rho_0/3)^{\frac{1}{3}} M^{5/3} + \cdots, \tag{15}$$

which in the first approximation is identical with (9). For large values of the proper mass $M \gg M_{\text{crit}}$ it is approximately

$$m = \alpha^{\frac{1}{2}} M^{\frac{1}{3}}. \tag{16}$$

According to (14) the effective mass m of a sphere of constant density therefore never becomes negative, but increases somewhat more slowly than M itself. Curve II in Fig. 1 shows the dependence of m on M according to the relation (14).

From what we have said already, it is clear that the classical theory of gravitation, even if combined with the fundamental results of the special theory of relativity is unsatisfactory, because it does not allow us, to formulate clearly how the mass equivalent of the gravitational potential energy is to be distributed over space. This difficulty has its real root in the fact that gravitation is a *cooperative phenomenon*.[4] Systems which are subject to the actions of such cooperative phenomena can be treated satisfactorily only with the help of theories which make possible the treatment of these systems in their entirety. The general theory of relativity is the only theory which permits a unified treatment of gravitation as a cooperative phenomenon. We therefore must now consider the results of this theory regarding the effective masses of large dense stars.

[4] F. Zwicky, Phys. Rev. **43**, 270 (1933).

C. The Schwarzschild Configuration

We again discuss the case of a sphere of constant density ρ_0. The general relativistic treatment of this case was first given by Schwarzschild.[5] If we neglect the cosmological constant the relativistic line elements for the exterior and the interior of the sphere have the following form:

exterior,

$$ds^2=-dr^2/(1-2\bar{m}/r)-r^2d\theta^2 -r^2\sin^2\theta d\phi^2+(1-2\bar{m}/r)dt^2; \quad (17)$$

interior,

$$ds^2=-dr^2/(1-r^2/R^2)-r^2d\theta^2 -r^2\sin^2\theta d\phi^2+[A-B(1-r^2/R^2)^{\frac{1}{2}}]^2dt^2; \quad (18)$$

where r, θ, ϕ and t are the polar coordinates and the time. The constants involved have the following values

$$A=\tfrac{3}{2}(1-r_1^2/R^2)^{\frac{1}{2}}, \quad B=\tfrac{1}{2}, \quad (19)$$

$$R^2=3/8\pi\bar{\rho}_0, \qquad \bar{m}=4\pi\bar{\rho}_0r_1^3/3, \quad (20)$$

where r_1 is the radius of our sphere, m is the effective mass of the sphere, which, in distant points ($r\gg r_1$), determines the gravitational field in the ordinary Newtonian way. The bars on $\bar{m}$, $\bar{M}$ and $\bar{\rho}_0$ indicate that the masses involved are expressed in the customary relativistic units rather than in ordinary units. These units for mass are so chosen that the mass m expressed in ordinary units is in the following manner related to the mass $\bar{m}$ which appears in the preceding relations.

$$\bar{m}=m\Gamma/c^2, \quad (21)$$

where

$$\Gamma/c^2=7.42\times10^{-29}\ \text{cm/g}. \quad (22)$$

If we assume that $\bar{\rho}_0$ is a constant throughout the sphere, its proper mass $\bar{M}$ is given by

$$\bar{M}=4\pi\bar{\rho}_0\int_0^{r_1}r^2(1-r^2/R^2)^{-\frac{1}{2}}dr. \quad (23)$$

Equations (20) and (23) determine the values of r_1 and $\bar{m}$ in terms of the proper mass $\bar{M}$ and the proper density $\bar{\rho}_0$. We rewrite these equations by introducing a new variable x, where

$$x=r_1/R. \quad (24)$$

Eq. (23) now reads

$$\bar{M}=4\pi R^3\bar{\rho}_0\int_0^x\xi^2(1-\xi^2)^{-\frac{1}{2}}d\xi \quad (25)$$

or $$\bar{M}/\bar{m}=(3/2x^3)[-x(1-x^2)^{\frac{1}{2}}+\arcsin x]. \quad (26)$$

From (26) it follows that there exists a mathematical limit[6] to the Schwarzschild solution (18) for $x=1$, which is equivalent to the following relations:

$$r_1=R, \quad r_1^2=3/8\pi\bar{\rho}_0, \quad r_1=2\bar{m}. \quad (27)$$

The physical significance of such a limit would be that the exterior space of the sphere would entirely disappear, and the interior of the sphere would represent the whole space available.

In addition to the fact that the Schwarzschild solution (18) breaks down already for $x=(8/9)^{\frac{1}{2}}$ it must be remarked that in actual space generally any complete closure of a star cannot take place for the following reason. In the exterior space of the sphere the average density ρ, although it is small, is nevertheless not exactly zero as we have assumed it. The exterior space can therefore not, so to speak, be entirely swallowed up by our sphere. In fact, even if ρ were initially equal to zero, it would necessarily cease to be equal to zero during the process of formation of the limiting Schwarzschild sphere. During this process the gravitational "packing" energy

$$|E_g|=(M-m)c^2 \quad (28)$$

[5] The formulation of the Schwarzschild solution used here may be found in R. C. Tolman, *Relativity, Thermodynamics, Cosmology* (Oxford, 1934).

[6] Actually this limit cannot be realized, because the Schwarzschild solution already breaks down at $x=(8/9)^{\frac{1}{2}}$, for which value of x the pressure becomes infinite at the center of the sphere. In the following section we shall nevertheless proceed to discuss the behavior in the whole range $0\leqslant x\leqslant 1$ of those properties of our star which do not exhibit any singularity at $x=(8/9)^{\frac{1}{2}}$. This facilitates the mathematical discussion since for $x=1$ the formulae involved take on a simpler form than if we inserted $x=(8/9)^{\frac{1}{2}}$. It must be remembered, however, that $x=(8/9)^{\frac{1}{2}}$ represents the maximum value of x for which any physical significance can be ascribed to the solution for a sphere of constant density which we have used here. In Section *G* we shall later on calculate some of the characteristics which interest us for the actual physical limit $x=(8/9)^{\frac{1}{2}}$. In this connection attention should be called to the fact that the Schwarzschild interior solution (18) does not seem to be the most general solution for a sphere of constant density. A discussion of a more general solution has just been published by Mr. G. M. Volkoff (Phys. Rev. 55, 413 (1939).

must be transferred from the sphere into its exterior, such that the exterior of the sphere necessarily will contain some energy or mass.

Limiting cases for which $\rho_0 \neq 0$ inside of the sphere and $\rho_e = 0$ outside, therefore would necessitate a rediscussion on the basis of a somewhat more refined picture regarding the distribution of mass over the various parts of the space considered. For instance as a next approximation a sphere of constant density ρ_0 imbedded in a universe of average density $\rho_e \ll \rho_0$ might be investigated. Nevertheless, as long as x is sensibly smaller than unity the corrections thus introduced will be negligible, and for small values of x the solution (17) and (18) as given by Schwarzschild will remain applicable to real cases. The discussion of the Schwarzschild formulae for $x<1$ leads to the following results.

As we have already stated, the expression $(M-m)c^2$ may be regarded as the energy which is released by the sphere into exterior space during the process in which the mass M dispersed over a large volume is condensed into the sphere of radius r_1. It is perhaps useful to show that for $x \ll 1$ the Schwarzschild expression for E_g is in the first approximation identical with the ordinary Newtonian value (1) of the gravitational packing energy of a sphere of constant density. Indeed, if we expand (26) in terms of x we obtain

$$\bar{M}/\bar{m} = M/m = 1 + 3x^2/10 + 9x^4/56 + \cdots. \quad (29)$$

From (20) and (24) we derive the explicit value of x in terms of ρ_0 and m. It is

$$x = (\pi/3)^{1/6} 2^{5/6} \tilde{\rho}_0^{1/6} \bar{m}^{\frac{1}{3}}, \quad (30)$$

or in terms of $\bar{m}$ and r_1

$$x^2 = 2\bar{m}/r_1. \quad (31)$$

If we return again to ordinary units of mass, (31) is equivalent to

$$x^2 = 2\Gamma m/r_1 c^2. \quad (32)$$

Therefore, in the first approximation

$$M/m = 1 + 3\Gamma m/5 r_1 c^2. \quad (33)$$

Since for $x \ll 1$ we have $M - m = \Delta \ll M$, we finally obtain

$$m = M - 3\Gamma M^2/5 r_1 c^2, \quad (34)$$

which is identical with the expression (9) for the effective mass of a sphere of constant density. Agreement, for weak gravitational fields, with the Newtonian theory is a necessary prerequisite of any general theory of gravitation. The general relativistic expressions (28) and (26) for the gravitational energy of spheres of uniform density are therefore in so far satisfactory as they converge towards the Newtonian expression (9) in the case of proper masses M which are small compared with the critical mass M_{crit} (6).

We now discuss somewhat more fully the dependence of the effective mass m in the Schwarzschild solution on the proper mass M. It can easily be shown that the difference between the Newtonian expression and the Schwarzschild expression for the gravitational E_g is inappreciable for all possible values of x. The real difference between the two expressions is not a numerical one. The essential difference rather consists in the fact that in the Newtonian case we may, at constant density ρ_0, increase the proper mass M of our sphere indefinitely without introducing any infinite values for the pressure or making the velocity of light anywhere equal to zero. In the general theory of relativity this is not possible because of the existence of critical masses some characteristics of which we shall discuss below.

For values $x>1$ the Schwarzschild solution breaks down mathematically since the expression $(1-x^2)^{\frac{1}{2}}$ which occurs in the line element becomes imaginary. For the mathematical real limit $x=1$ we obtain from (26)

$$\bar{M}_L/\bar{m}_L = M_L/m_L = 3\pi/4. \quad (35)$$

This means that the gravitational energy E_{gL} which is liberated in the process of contraction of matter of the proper mass M_L from the completely dispersed state into the configuration $x=1$ is equal to

$$-E_{gL} = (M_L - m_L)c^2 = M_L(1-4/3\pi)c^2 \quad (36)$$

or

$$-E_{gL} = 0.58 M_L c^2. \quad (37)$$

Consequently somewhat more than half the proper mass would be transformed into energy during the formation of the sphere $x=1$.

The relation between M_L and the density ρ_0 follows from (27) and (35) with the result that

$$M_L^2 \rho_0 = \pi(3c^2/8\Gamma)^3. \quad (38)$$

The Newtonian expression for the gravitational

energy E_{gL}' of a sphere of mass M_L and of constant density ρ_0 is, according to (3)

$$E_{gL}' = -\tfrac{3}{5}(4\pi/3)^{\frac{1}{3}}\Gamma M_L^{5/3}\rho_0^{\frac{1}{3}} \tag{39}$$

or with substitution of (38)

$$E_{gL}' = -(9/40)(4/3)^{\frac{1}{3}}\pi^{\frac{1}{3}}M_L c^2 = -0.53 M_L c^2. \tag{40}$$

The difference between the general relativistic and the Newtonian values of the gravitational energy for $M = M_L$, therefore, is

$$E_{gL} - E_{gL}' = -0.05 M_L c^2, \tag{41}$$

which is less than ten percent of the respective values of E. This result is perhaps somewhat unexpected if we consider that for highly collapsed configurations the Schwarzschild solution presents conceptually and actually aspects so fundamentally at variance with the Newtonian theory. The differences between the two theories are especially striking when the properties of light signals from our sphere to a distant observer outside are investigated. We shall come back to this problem in Sections F and G.

We now show that the effective mass m of a sphere of constant density ρ_0 and of proper mass M as given by Schwarzschild's solution is for all possible values of ρ_0 and M slightly smaller than the effective mass m given by the Newtonian expression (9) or the corrected Newtonian expression (14). For this purpose we investigate the respective values of the derivatives $(dm/dM)_{\rho_0=\text{const}}$. From the Schwarzschild solution (26) we derive

$$(dM/dm)_{\rho_0=\text{const}} = \frac{1}{2}\left[\frac{3}{x^3} - \frac{9m}{x^4}\frac{dx}{dm}\right] \times [-x(1-x^2)^{\frac{1}{2}} + \arcsin x] + \frac{3m}{2x^3}\left[-(1-x^2)^{\frac{1}{2}} + \frac{x^2+1}{(1-x^2)^{\frac{1}{2}}}\right]\frac{dx}{dm}. \tag{42}$$

From (30) it follows

$$dx/dm = x/3m. \tag{43}$$

Substituting (43) in (42) we obtain

$$(dm/dM)_{\rho_0=\text{const}} = (1-x^2)^{\frac{1}{2}} \tag{44}$$

or with (31) and (44)

$$(dm/dM)_{\rho_0=\text{const}} = [1 - 2\Gamma(4\pi\rho_0/3)^{\frac{1}{3}} m^{\frac{2}{3}}/c^2]^{\frac{1}{2}}. \tag{45}$$

On the other hand, we derive from the Newtonian expression (9) the following relation

$$(dm/dM)'_{\rho_0=\text{const}} = 1 - \Gamma(4\pi\rho_0/3)^{\frac{1}{3}} M^{\frac{2}{3}}/c^2. \tag{46}$$

Since always $m < M$ it is seen that for all values of M and ρ_0

$$(dm/dM)' > (dm/dM) \tag{47}$$

or, in turn

$$m(\rho_0, M)_{\text{Newton}} > m(\rho_0, M)_{\text{Schwarzschild}}. \tag{48}$$

Again for values of $M \ll M_{\text{crit}}$ it is approximately

$$(dm/dM)'_{\text{Newton}} = (dm/dM)_{\text{Schwarzschild}}, \tag{49}$$

which is identical with our former conclusion (34) that for small proper masses M the general theory of relativity and the Newtonian theory lead to the same gravitational energy of our sphere. The curve III in Fig. 1 represents $m(M)$ of a sphere of constant density $\rho_0 = 10^{14}$ g/cm^3 as given by the Schwarzschild solution.

D. The Schwarzschild Mass M_L for Neutron Stars

According to (35) it would be, for $x = 1$

$$m_L = 4M_L/3\pi \tag{50}$$

and from (30)

$$x = (\pi/3)^{1/6} 2^{5/6} \rho_0^{1/6} m_L^{\frac{1}{3}} (\Gamma/c^2)^{\frac{1}{2}} = 1. \tag{51}$$

Substituting (50) in (51) we obtain

$$M_L = 2^{-5/2} 3^{\frac{1}{2}} \pi^{\frac{1}{2}} (\Gamma/c^2)^{-\frac{3}{2}} \rho_0^{-\frac{1}{2}} \tag{52}$$

or numerically

$$M_L = 6.39 \times 10^{41} \rho_0^{-\frac{1}{2}}. \tag{53}$$

If we adopt $\rho_0 = 10^{14}$ g/cm^3 as a standard nuclear density which is of the order of the density to be expected in collapsed neutron stars, the characteristic proper mass M_L becomes[7]

$$M_L = 6.4 \times 10^{34} \text{ g}. \tag{54}$$

This mass is approximately equal to $32 M(\odot)$.

[7] Following the discussion given by R. C. Tolman in *Relativity, Thermodynamics and Cosmology* (Oxford, 1934), p. 247, I have recently stated (F. Zwicky, Astrophys. J. 88, 522 (1938)) that M_L represents a natural upper limit for the mass of neutron stars inasmuch as for this mass a neutron star would entirely close itself and thus represent the whole space available. In view of the reasons given in Section E such an extreme configuration can, however, probably never be realized.

In regard to some views recently advanced[8] concerning possible changes of dimensionless ratios among the so-called fundamental physical constants it is of interest to express M_L or M_l (see Section G) in terms of these constants. This may be done in the following way. Instead of adopting for ρ_0 a definite numerical value we express ρ_0 approximately in terms of fundamental physical constants. We start with the expression for a length δ of nuclear dimension

$$\delta = e^2/m_e c^2 = 2.81 \times 10^{-13} \text{ cm}, \tag{55}$$

where e and m_e are the charge and mass of an electron. As the prototype v of a nuclear volume we may take the sphere of radius δ such that

$$v = 4\pi\delta^3/3 = 9.3 \times 10^{-38} \text{ cm}^3. \tag{56}$$

If the volume v contains the mass m_P of a proton (or a neutron), the average density ρ' in v is

$$\rho' = m_P/v = 1.78 \times 10^{13} \text{ g/cm}^3. \tag{57}$$

In an actual collapsed neutron star the volume occupied by each neutron is equal to γv where γ is of the order of unity (probably about equal to $\gamma = 5$). We therefore introduce in (52) the following expression for the density

$$\rho_0 = m_P/\gamma v = \rho'/\gamma \tag{58}$$

and we obtain for the prototype (p) of a highly collapsed star

$$M_L(p) = \gamma^{\frac{1}{2}} 3\pi 2^{-7/2} (e^2/\Gamma m_P m_e)^{\frac{1}{2}} m_N \tag{59}$$

or

$$M_L(p) = 0.83 \gamma^{\frac{1}{2}} R^{\frac{1}{2}} m_N, \tag{60}$$

where $R = 2.29 \times 10^{39}$ is the well-known dimensionless number which represents the ratio between the electric and the gravitational attraction between an electron and a proton.

For an actual neutron star the constant γ may be written as a product

$$\gamma = \gamma_1 \gamma_2, \tag{61}$$

where $\gamma_2 \neq 1$ indicates that the neutron, considered as a hard sphere has a proper density $\rho = \rho'/\gamma_2$, whereas $\gamma_1 \neq 1$ indicates that if we try to pack hard spheres as closely as possible there still remain interstices which amount to 26 percent of the total volume. Consequently for an assembly of closely packed spheres

$$\gamma_1 = 100/(100-26) = 1.35.$$

Concerning the numerical value of γ_2 we may make no very precise statement. Since many neutrons are packed into ordinary heavy nuclei we might adopt the average nuclear density of such nuclei as a reasonable value for ρ_0, as we have done in calculating M_L as given by (54). It must be remembered, however, that some uncertainty arises regarding the actual value of γ_2 and therefore of ρ_0 because of the tremendous gravitational pressure in neutron stars. This pressure affects the closest approach or the packing of neutrons in a way which cannot be predicted without any further studies of the fundamental properties of neutrons. Also it must be remembered that a neutron star represents a Fermi degenerate state with resulting kinetic energies of the individual neutrons which may be quite comparable with the proper energy $m_N c^2$ of a neutron. The next step of the theoretical approach to neutron stars therefore must take into account the *equation of state* of a degenerate neutron star. This, however, lies beyond the scope of the preliminary outline of the theory of highly collapsed stars given in this paper.

E. Relation to Some Cosmological Speculations

Finally a few remarks concerning the possible *cosmological* importance of the dimensionless ratio R may be in order. Many years ago H. Weyl[9] first remarked that the tremendous value of R suggests a possible connection of R with other equally large numbers. Astronomical observations during the last two decades have brought to light the existence of several numbers of this kind. These observations refer in particular to the redshift of light from distant nebulae and to the distribution of nebulae in space.

Denote with ν the frequency of a specific spectral line, such as the H or K line of calcium, as measured for the light of a terrestrial source. The frequency ν' of the same spectral line in the

[8] F. Zwicky, Phys. Rev. **43**, 1031 (1933) and **53**, 315 (1938); Philosophy of Science **1**, 353 (1934); Proc. Nat. Acad. Sci. **23**, 106–110 (1937).

[9] H. Weyl, *Raum. Zeits. Materie*, 5th edition (Springer, Berlin, 1923), p. 277. See also Naturwiss. **22**, 145 (1934).

light which reaches the earth from a distant nebula in the first approximation is $\nu' = \nu + \Delta\nu$, where

$$\Delta\nu/\nu = -D/R'. \tag{63}$$

If the distance D of the nebula is measured in units of $\delta = e^2/m_e c^2$, then R' according to Hubble and Humason is a dimensionless number of the order

$$R' = 6 \times 10^{39}. \tag{64}$$

A second large number N whose order of magnitude may be determined by direct observations is the number of elementary particles, such as protons and electrons, which is contained in the sphere the radius of which is $R'\delta$, where

$$R'\delta = 1.68 \times 10^{27} \text{ cm} \cong 1.7 \times 10^{9} \text{ light years}. \tag{65}$$

If ρ_s is the average density of matter in this sphere we have

$$N \cong (4\pi/3)(R'\delta)^3 \rho_s / m_P. \tag{66}$$

From the results of counts of nebulae, combined with an estimate of the mass of average nebulae, we may assume in order of magnitude

$$\rho_s \cong 10^{-28} \text{ g/cm}^3. \tag{67}$$

Consequently $$N \cong 1.2 \times 10^{78}, \tag{68}$$

which is not far different from R^2.

A third large number R'' might ultimately be derived from observations resulting in a possible curvature of space. Such observations have been attempted by Hubble who counted nebulae to the limit reached by the 100-inch telescope. The apparent too rapid increase in the number of nebulae per unit volume at very great distances might be interpreted as indicating the existence of a curvature of space which is different from zero. Since certain objections can be raised against such an interpretation, we shall for the present not longer dwell on the discussion of R''.

As was already stated, Weyl[9] and more recently Dirac[10] and others have suggested that there exist relations among the dimensionless numbers R, R', R'' and N. Such an hypothesis must naturally be regarded as mere speculation unless it allows us to draw further conclusions which are in agreement with observational facts. Such conclusions can actually be drawn, as I have previously pointed out,[8] if we examine Weyl's hypothesis in the light of the fundamental question, whether or not the average physical conditions in our universe are slowly but systematically changing. For instance, if the redshift of light from distant nebulae is to be interpreted as an average velocity of recession of all nebulae from any observer fixed to a given nebula, we should be confronted with such a systematic change (expansion of the universe). On the other hand, if the redshift should find another explanation leaving the dimensions of the universe stationary in the sense that R' is an absolute constant we might find that matter on the average is transforming itself gradually into radiation. In any of these cases we may be forced to conclude that either R', or N, or both are changing. If invariant relations exist between R, R', R'' and N it follows that all of these numbers change if any one among them changes. Such a conclusion, however, leads to the possibility of further tests of Weyl's hypothesis. One such possibility is indicated by the following line of argument.

In the sense of Sections D and G we may regard the expression $R^1 m_N$ as the prototype of a stellar mass. Since, according to the mass-luminosity relation the brightness of a star is essentially determined by its mass, a change in R should result in a change of the average brightness of stars. In the specific case that R' and R slowly change their values because of a mutual recession of all nebulae this leads to a change in the average brightness of stars and nebulae as a function of the distance D from the observer. Numbers of nebulae to given successively fainter limiting magnitudes in this way would be theoretically related to the magnitude of the redshift observed on the same limiting nebulae. This conclusion from Weyl's hypothesis therefore is subject to direct tests through counts of nebulae to ever decreasing apparent brightness coupled with observations on the redshift. It is important to notice that in such a test the very difficult determination of the distances of faint nebulae is entirely dispensed with.

We briefly mention one other cosmological aspect of the theory of neutron stars. O. Stern[11]

[10] P. A. M. Dirac, Nature **139**, 323 (1937).

[11] O. Stern, Zeits. f. Elektrochemie **31**, 448 (1925).

first pointed out that the application of the principles of statistical mechanics and of thermodynamics to the equilibrium between matter and radiation in the universe leads to the conclusion that in a state of equilibrium practically all matter would be found to have been transformed into radiation. The fact that stars are radiating energy continuously into empty space, combined with our present ignorance of any phenomena which transform this radiation back into matter might be taken as a confirmation of Stern's statement that in the progression of the universe towards a stationary state the systematic transformation of matter into radiation would seem inevitable. The problem raised by Stern was subsequently discussed by many investigators[12] without any essential difference in the final result being achieved. All attempts to save matter from its final annihilation without abandoning entirely the principles of statistical mechanics were blocked by the fact that according to these principles the ratio of the concentration c_m of matter in the ordinary forms and of the concentration c_R of matter transformed into radiation were determined by a Boltzmann factor of undefeatable magnitude, namely

$$c_R/c_m \sim e^{mc^2/KT}. \tag{69}$$

Here m is the mass of an elementary particle, c the velocity of light, $K=1.37\times10^{-16}$ erg Boltzmann's constant and T the average absolute temperature in the universe. Inserting for m the mass of the electron m_e and assuming for T as high a value as 100° abs. we find that the Boltzmann factor in (62) is still equal to about

$$e^{m_e c^2/100K} \sim e^{6\times10^7}, \tag{70}$$

which means that in a sphere of the enormous radius R' not even a single electron could exist in statistical competition with radiation. As a possible escape from this unpleasant conclusion I pointed out some time ago[13] that it is not permissible to consider only matter in its ordinary free state in the preceding discussion. Matter may indeed assume states of high statistical weight through gravitational clustering. If this clustering leads to configurations, as it does in limiting neutron stars, where the velocity of light becomes zero, the gravitational packing energy of a particle of mass m may become equal to its original proper energy mc^2. Such configurations of matter consequently are distinguished by Boltzmann factors of precisely the same magnitude as (69) and matter in this form may successfully enter into statistical competition with radiation. It is obvious, however, that a satisfactory solution of the whole problem must also include an estimate not only of the Boltzmann factors but also of the so-called *a priori* weights of matter and of radiation.

[12] R. C. Tolman, Proc. Nat. Acad. Sci. **12**, 670 (1926).
[13] F. Zwicky, Proc. Nat. Acad. Sci. **14**, 592 (1928).

F. Velocity and Redshift of Light Emitted from the Surface of the Star

A light signal which travels radially from the star is described by the conditions

$$ds=0, \qquad d\theta=d\phi=0. \tag{71}$$

In combination with the expression (17) for the line element in the exterior of the sphere this leads to the following value c' for the velocity of light

$$\overline{c'}=dr/dt=1-2\bar{m}/r \tag{72}$$

or if time t and mass m are measured in ordinary units

$$c'=dr/dt=c(1-2\Gamma m/rc^2), \tag{73}$$

where c' and c are also in ordinary units. For $x=1$ it would be $2\bar{m}/r_1=1$ on the surface of the star. Consequently $c'=0$ on the surface. For values of x close to $x=1$ the velocity of light will be almost equal to zero for $r=r_1$ and will gradually grow towards its ordinary value c for $r\gg r_1$. A light signal, for its journey from the surface of the star to an external point $r>r_1$ therefore consumes a time which becomes longer and longer as x grows and which approaches an infinite value as x approaches unity. For the reasons discussed in Sections C and G these extreme conditions are, however, unlikely to be ever actually realized.

The redshift of light from the surface of the star is given by the following relation

$$\lambda'/\lambda=dt/ds=(1-2\bar{m}/r)^{-\frac{1}{2}}. \tag{74}$$

Here λ is the wave-length of a spectral line observed in a point P far away from the star which an atom would emit if placed itself in or near

the point P. On the other hand, λ' is the wave-length of the same spectral line again observed in P but emitted by an identical atom placed at the distance r from the center of the star.

If we approach $x=1$ we see that for r approaching r_1 the wave-length λ' grows ever larger and for $x=1$ would become infinite.

The results expressed in (63) and (64) signify that the physical communication between a point P_s on the surface of our collapsed star and a point outside become increasingly difficult as the mass and the density of a star increase, both because the light traveling from P_s to P consumes an ever increasing time and because quanta of light originating in familiar emission processes reach the point P with energies which appear more and more depleted as compared with the energy of quanta produced by the identical processes of emission in the point P itself.

From (73) and (74) the velocity of light c' on the surface of a sphere of constant density and the redshift of light from this surface, for the characteristic value $x=(8/9)^{\frac{1}{2}}$, are

$$c'=c/9, \tag{73a}$$

and

$$(\lambda'/\lambda)_l=3. \tag{74a}$$

What the extreme physically limiting values of c' and λ'/λ for spheres of arbitrary density distribution will turn out to be depends on the solution of the variation problem discussed in Section H, the solution of which would acquaint us with the properties of ultimately collapsed configurations of matter as such.

G. Characteristics of the Schwarzschild Solution for $x=(8/9)^{\frac{1}{2}}$

In Section C (reference 6) we stated that Schwarzschild's solution breaks down already for certain values of $x<1$. This is due to the fact that in the center of our sphere the pressure becomes infinite and the velocity of light equal to zero when $x=(8/9)^{\frac{1}{2}}$ as Eddington[14] has pointed out. The expression for the pressure p in the interior of the sphere is

$$8\pi p=R^{-2}[3B(1-r^2/R^2)^{\frac{1}{2}}-A]/[A-B(1-r^2/R^2)^{\frac{1}{2}}], \tag{75}$$

where A and B are given by (19). The pressure therefore becomes infinite, if

$$A-B(1-r^2/R^2)^{\frac{1}{2}}=0 \tag{76}$$

or

$$r^2=9r_1^2-8R^2. \tag{77}$$

As long as $8R^2>9r_1^2$ Schwarzschild's solution is satisfactory, inasmuch as there exists no real value of r for which p becomes infinite. For $8R^2\leq 9r_1^2$ the radius r of the sphere on whose surface the pressure becomes infinite assumes real values. The range of validity of Schwarzschild's solution (18) therefore is described by

$$0\leq x=r_1/R\leq(8/9)^{\frac{1}{2}}. \tag{78}$$

From (19) it is

$$A=B=\tfrac{1}{2} \quad \text{for} \quad x=(8/9)^{\frac{1}{2}}. \tag{79}$$

The velocity c'' of light in the interior of our sphere, from Eq. (18), in ordinary units of time, is

$$c''=c[A-B(1-r^2/R^2)^{\frac{1}{2}}](1-r^2/R^2)^{\frac{1}{2}} \tag{72a}$$

or for $x=(8/9)^{\frac{1}{2}}$, with substitution of (79)

$$c''=0 \quad \text{for} \quad r=0. \tag{73b}$$

With $x=(8/9)^{\frac{1}{2}}$, we obtain the following characteristic values for the proper mass M_l and the effective mass m_l of a sphere of constant density ρ_0

$$m_l=0.61M_l \tag{80}$$

and

$$M_l=3.75\times 10^{41}\rho_0^{-\frac{1}{2}}. \tag{81}$$

The gravitational packing energy E_{gl} for the sphere of constant density ρ_0 and of proper mass M_l consequently is

$$-E_{gl}=(M_l-m_l)c^2=0.39M_lc^2. \tag{82}$$

For a neutron star with a density of the order $\rho_0=10^{14}$ g/cm^3 it would be

$$M_l=3.75\times 10^{34}\,\text{g}\cong 19M(\odot) \tag{83}$$

$$m_l=2.26\times 10^{34}\,\text{g} \tag{84}$$

and from (44)

$$(dm/dM)_l=\tfrac{1}{3}. \tag{85}$$

[14] A. S. Eddington, *The Mathematical Theory of Relativity* (Cambridge, 1923), p. 169.

H. Further Problems Concerning Collapsed Neutron Stars

From the preceding considerations the following mathematical and physical problems suggest themselves.

We are obviously interested in finding that configuration of matter of the proper mass M which possesses the lowest possible effective mass m. A body whose m/M is an *absolute minimum* represents the configuration of lowest energy which matter may assume without being completely transformed into radiation. If we assume for the moment that the general theory of relativity will supply the correct answer, the problem to determine this configuration of lowest energy is equivalent, mathematically, to a problem of the calculus of variations. The fact that the condition $p=\infty$ determines the proper mass M_1 for which the velocity of light becomes zero at the center of our sphere suggests that we must, for $M=$const. and $p=\infty$ at the center of the sphere, find that distribution $\rho_0(r)$ of the proper density as a function of r which will make the effective mass m an absolute minimum. The limiting condition $p=\infty$ might possibly have to be replaced by the velocity of light $c=0$, if the two conditions are not equivalent.

Physically it might, however, happen that a configuration of minimum m cannot be formed in a finite time. For instance in the case of the characteristic Schwarzschild sphere ($M=M_1$) the velocity of light becomes equal to zero at the center and with it the velocities of all physical reactions tend towards zero. Since the corresponding velocities in points $r\neq 0$ do not become zero simultaneously, the theory of the process of formation of the characteristic sphere M_1 must necessarily consider nonstatic configurations which are caused by the queer nonisotropic behavior of the time scale.

Finally we must not forget that actual matter is subject to restrictions expressed commonly by *equations of state*. For instance, matter in the highly collapsed state is subject to some type of quantum degeneracy as has been pointed out by a number of investigators.[15] The ultimate problem before us therefore is to find the configuration of absolutely lowest energy which at the same time satisfies the equilibrium conditions demanded by the actual equations of state of matter in highly collapsed stars. No attempt is here made to determine this configuration.

I. Neutron Stars and Actual Stars

The following problems suggest themselves. (a) Do stars exist which are configurations of lowest energy or at least configurations approaching this limit? (b) Do any of the ordinary stars possess cores of highly collapsed matter of the type described in the preceding? (c) How is the transformation from stars composed of ordinary matter into highly collapsed stars accomplished?

We shall not here give any discussion of the first two questions. The third question, however, may find its answer through observations of supernovae. Some of the reasons which make the formation of neutron stars as a cause for supernovae an attractive possibility are roughly as follows.

(1) On the neutron star hypothesis it is easy to account for the tremendous liberation of energy which we have observed in the supernova outbursts.

(2) A second characteristic of supernovae is that the transformation from a given stellar configuration into an apparently entirely different configuration takes place at a fantastic speed as compared with the slow changes in the characteristics of ordinary stars. The explanation of the explosive character of supernovae may be related to the fact, unique in neutron stars, that such stars possess a nearly vanishingly small opacity.

(3) The spectra of supernovae are entirely different from those of any known stellar spectra. It will become clear from the following discussion, that on the neutron star hypothesis the physical conditions which determine the character of the emission and absorption of light on the effective surface of supernovae are indeed so unusual that the resulting spectra must be entirely different from anything observed so far.

(4) The progressive redshift of all of the *permanent* features (blue bands) of the spectra of supernovae found first by R. Minkowski[16]

[15] G. Gamow, *Atomic Nuclei* (Oxford, 1937), p. 234; L. Landau, Nature **141**, 334 (1938).

[16] R. Minkowski, Astrophys. J. **89** 156 (1939); F. Zwicky *ibid.* **88**, 522 (1938).

must at the present be regarded as one of the most important features. On the neutron star hypothesis this redshift is readily explainable as a general relativistic gravitational redshift.

(5) For a common nova the difference $m_f - m_{max}$ between the photographic magnitudes of its final stage and its maximum brightness apparently never exceeds fifteen magnitudes. Although we do not, as yet, possess any data regarding the final stages (remnants) of supernovae good evidence is at hand that the *stellar remnants* of supernovae in some cases must be fainter than the supernovae at maximum by at least twenty magnitudes. On the basis of the neutron star hypothesis this may find its explanation in the fact mentioned in (3) that neutron stars in themselves are incapable of retaining any excessive radiation density.

Before we discuss some of the preceding suggestions in more detail a rough estimate may be made regarding the critical value of stellar masses for which the formation of neutron stars is energetically more favorable than the transformation from one stellar configuration into another which is caused by any known type of nuclear reaction.

We may assume in order of magnitude that the most efficient nuclear reaction which can take place throughout the whole mass M of a given star liberates an energy

$$E \leqq 10^{-3} Mc^2. \tag{86}$$

According to (4) the energy E_g liberated through the formation of a neutron star of density $\rho_0 = 10^{14}\,\mathrm{g/cm^3}$ becomes greater than E for masses M which satisfy the inequality

$$M > M_0 = 5.2 \times 10^{30}\,\mathrm{g} \leqq M(\odot)/400. \tag{87}$$

This condition is therefore fulfilled by most of the ordinary stars.

K. Redshift in the Spectrum of Supernovae

The spectra of about half a dozen supernovae have been observed thus far. Photographic records over periods of many months, however, have been secured only for the recent supernovae in IC 4182 and NGC 1003. A detailed account of these observations was published by R. Minkowski.[16] He finds that the spectra of all supernovae observed so far are surprisingly similar and that they exhibit a number of wide emission bands in the red and in the blue, respectively. Whereas the red bands in course of time show considerable changes in relative intensity and finally, about nine months after maximum brightness fade away entirely, the blue bands seem to be of a more *permanent* character, except for the fact that their respective intensity maxima are subject to a gradual shift towards longer wavelengths. The total shift of the blue bands towards the red end of the spectrum amounts to about 100A in the period of eight months. The average rate of change of the red shift during a few months following the date of maximum brightness is essentially the *same* for the supernovae which appeared in the spirals NGC 4273, 1003 and IC 4182 regardless of the fact that the absolute brightness of these supernovae at maximum was different by as much as a factor forty.

To fix our ideas, the processes which lead to a progressive redshift of the blue emission bands may schematically be pictured in the following manner. There exists a characteristic surface of separation Σ_s in supernovae[2] such that the material outside this surface is expelled into interstellar space, while matter inside of Σ_s gradually contracts. It is this matter inside of Σ_s which constitutes the stellar remnant of the supernova on whose surface the blue emission bands originate. While this stellar remnant contracts, the gravitational field on its surface becomes increasingly stronger and causes the frequency of characteristic spectral lines to diminish continuously.

For the following considerations we shall make use of the data[16, 17] obtained for the absolutely brightest of all supernovae which are known at the present time. This supernova flared up in the faint spiral IC 4182. It presumably reached its maximum brightness on August 22, 1937, at which time its absolute photographic magnitude was $\mathfrak{M}_p = -16.6$ corresponding to six hundred million times the radiation of the sun in the photographic spectral range. Although we have not yet accurately determined the radiation of this supernova in the red parts of the spectrum it follows from our data that at maximum its

[17] W. Baade and F. Zwicky, Astrophys. J. 88, 411 (1938).

total radiation (per second) L_{max} in the spectral range from $\lambda=7000$A to $\lambda=3300$A is approximately equal to a billion times that of the sun ($L(\odot)=3.8\times10^{33}$ ergs/sec.). We therefore have

$$L_{max}\leqq 3.8\times10^{42}\text{ ergs/sec.} \tag{87}$$

By the end of May, 1938 the absolute photographic brightness had become about $\mathfrak{M}_p=-9.5$. The total radiation in the visible range at this time therefore had decreased to approximately $L=L_{max}/1000$ or

$$L\leqq 3.8\times10^{39}\text{ ergs/sec.} \tag{88}$$

In the same period, from August 22, 1937 to May 30, 1938, the blue bands had shifted towards the red by an amount approximately equal to $\Delta\lambda=100$A. For purposes of calculation we shall set

$$\kappa=\Delta\lambda/\lambda=0.02. \tag{90}$$

In comparison we notice that the corresponding values of κ for the sun and the companion of Sirius are 2.1×10^{-6} and 6×10^{-5}, respectively.

On the basis of the neutron star hypothesis it is suggestive to interpret this redshift in supernovae as a gravitational redshift and to assume that the emission bands which are subject to this redshift originate on the surface of the central stellar remnant of the supernova. This is in keeping with the fact that these bands are much more permanent in character than the red bands whose pronounced changes in intensity combined with the absence of a redshift suggest that these red bands originate in more outlying layers of gas which have been expelled from the central star.

We now proceed to develop some further consequences of this hypothesis.

L. Resulting Radius of the Stellar Remnant of the Supernova in IC 4182

The change of the wave-length of light on its journey from the surface of a star of effective mass m and radius r_1 to a distant point $r\gg r_1$ is given by (74). We write for the wave-length λ' in the external point, in terms of the wave-length λ on the surface of the star

$$\lambda'=\lambda+\Delta\lambda, \tag{91}$$

where

$$\kappa=\Delta\lambda/\lambda=(1-2\Gamma m/r_1c^2)^{-\frac{1}{2}}-1. \tag{92}$$

Since we shall have to deal with small values of κ only, we may expand this expression into the one commonly used in discussions of stellar gravitational redshifts. It is

$$\kappa=\Gamma M/r_1c^2. \tag{93}$$

In conformity with the approximation used, the effective mass m has been replaced by the proper mass M. According to (93) and (90) we therefore obtain for the stellar remnant of the supernova IC 4182 (in the stage of June, 1938)

$$r_1=7.4\times10^{-29}M/\kappa=3.7\times10^{-27}M. \tag{94}$$

Offhand we do not know what the mass M of the stellar remnant of our supernova is and we therefore cannot immediately determine its radius, its density and some of its other physical characteristics. On the neutron star hypothesis it is, however, possible to estimate the maximum and minimum values of M.

As will be shown in the next section the minimum value M_{min} satisfies the inequality

$$M_{min}>2.16\times10^{32}\text{ g}=M(\odot)/9. \tag{95}$$

On the other hand, the neutron star hypothesis suggests that according to the considerations given in Section G we have a characteristic mass M_l. The conclusion that in a neutron star of this mass the velocity of light and with it the velocity of all reactions becomes zero at the center of the star might be taken as an indication that the characteristic mass M_l (83) represents a natural limit for the upper value M_u of the mass of the stellar remnant, such that

$$M_u<M_l=3.8\times10^{34}\text{ g}=19M(\odot). \tag{96}$$

In our special case we also possess some direct observational evidence which corroborates the correctness of (96). From the fact that the original star of the supernova in IC 4182 before its outburst was fainter than the apparent photographic magnitude 21, combined with our knowledge of the distance of IC 4182 (900,000 parsecs), we conclude that the absolute magnitude of the said original star must have been fainter than -4. It therefore was not at all one of the brightest stars in IC 4182, and its mass M_u, according to the mass luminosity relation, presumably was

$$M_u<15M(\odot)=3\times10^{34}\text{ g}. \tag{97}$$

In the case of supernovae in elliptical nebulae, such as the most recent supernova, which appeared in NGC 4636, we may even say that the original star did not have an absolute magnitude brighter than about -2. If, therefore, the redshift in supernovae is a gravitational redshift, Table I describes the range of possible radii of the stellar remnant of the supernova in IC 4182 and presumably of supernovae in general.

M. Density of the Stellar Remnant of Supernovae

According to the gravitational interpretation of the redshift in the spectrum of supernovae the average density ρ_a of their stellar remnants is given by

$$\rho_a = 3M/4\pi r_1^3 = 3(c^2/\Gamma)^3\kappa^3/4\pi M^2 \qquad (98)$$

or numerically, with the observed value $\kappa = 0.02$

$$\rho_a = 4.7\times 10^{78} M^{-2}. \qquad (99)$$

Table II gives the range of possible values of the average density ρ_a. From this table it follows that on our interpretation of the redshift as a general relativistic effect, combined with the neutron star hypothesis, masses smaller than $M_{\min}$ must be ruled out, since otherwise the introduction of densities surpassing those of neutron stars would be necessary. If it should be found that densities greater than 10^{14} g/cm^3 are possible, the value of $M_{\min}$ must be correspondingly decreased.

The actual range of possible masses is further narrowed down by the fact that the density cannot be assumed to be uniform throughout the stellar remnant of a supernova. A star whose density throughout the star is about 10^{14} g/cm^3 would be a *pure* neutron star. Such a star, because of its lack of opacity, could not retain any radiation longer than a few seconds, a conclusion which obviously contradicts the fact that the stellar remnant of the supernova in the stage which we are investigating (end of May, 1938) still was a million times as bright as the sun and at the time of writing (February, 1939) continues to radiate at a rate surpassing that of the sun by at least a factor 10,000. Here again attention must be called to the fact that our whole discussion is based on the assumption that the observed redshift is a gravitational redshift. If this assumption should prove to be incorrect, we could of course not conclude that the blue emission bands, which are subject to the redshift, have their origin on the surface of the stellar remnant of the supernova. These bands might then for instance originate in an extended gaseous envelope which presumably surrounds the stellar remnant of a supernova.

Table I. *Possible radii of the stellar remnant of the supernova in IC 4182. M is the proper mass of the central stellar remnant.*

M in g	r_1
2.2×10^{33}	8 km
$2\times 10^{33} = M(\odot)$	74 km $= 10^{-4} r(\odot)$
3×10^{34}	1100 km

Table II. *Possible values of the average density in supernovae.*

M in g	ρ_a in g/cm^3
$2.16\times 10^{32} (= M_{\min})$	10^{14}
$2\times 10^{33} (= M\odot)$	1.18×10^{12}
$3\times 10^{34} (= M_u)$	5.2×10^{9}

From these considerations we conclude that the possible mass M of the stellar remnant of the supernova IC 4182 must fall into the range

$$2.16\times 10^{32}\ \text{g} < M < 3\times 10^{34}\ \text{g}, \qquad (100)$$

a range which is practically identical with that covering all of the ordinary stars.

N. On the Total Generation of Energy in Supernovae

It was shown in Section C (41) that the relativistic and the Newtonian values for the gravitational energy of contraction E_g of a sphere of constant density differ at most by a few percent. Especially for small values of the energy, that is for $|E_g| \ll Mc^2$ we may always replace the relativistic value by the Newtonian value and write

$$E_g = -3\Gamma M^2/5r_1 = -6.44\times 10^{-8} M^{5/3}\rho_a^{1/3}. \qquad (101)$$

Substituting (93) in (101) we obtain

$$E_g = -3\kappa Mc^2/5, \qquad (102)$$

and for the difference between proper mass M and effective mass m

$$\Delta M = M - m = -E_g/c^2 = 3\kappa M/5. \qquad (103)$$

TABLE III. *Significant energy values for the supernova in IC 4182 as functions of possible masses M of its stellar remnant.*

M IN G	E_t IN ERGS	E_{MAX}	Mc^2	E_{VIS}
$2.16\times10^{33}(=M_{min})$	$=2.33\times10^{51}$	2.33×10^{51}	1.9×10^{53}	10^{49}
$2\times10^{33}(=M(\odot))$	$\geq2.16\times10^{52}$	9.52×10^{52}	1.8×10^{54}	10^{49}
$3\times10^{33}(=M_u)$	$\geq3.24\times10^{53}$	8.60×10^{54}	2.7×10^{55}	10^{49}
$3.8\times10^{33}(=M_l)$	$\geq4.08\times10^{53}$	1.31×10^{55}	3.4×10^{55}	10^{49}

For the stellar remnant of the supernova IC 4182, with $\kappa=0.02$ it is

$$\Delta M=0.012M. \tag{104}$$

For $M=M_{min}=2.16\times10^{33}$ g the energy of contraction released by the supernova IC 4182 from August, 1937 until June, 1938 would therefore be

$$-(E_g)_{min}=\Delta Mc^2=2.33\times10^{51}\text{ ergs}. \tag{105}$$

This energy, for $M=M_{min}$ is identical with the total energy which is released through the contraction of the initial star into a pure neutron star. As we have previously stated it is actually $M>M_{min}$ and the total energy E_t released by the supernova in the stated period must therefore be greater than $-(E_g)_{min}$. Consequently we have

$$E_t>2.33\times10^{51}\text{ ergs}. \tag{106}$$

For comparison we add

$$E_{vis}=10^{49}\text{ ergs} \tag{107}$$

and

$$Mc^2>1.9\times10^{53}\text{ ergs} \tag{108}$$

where E_{vis} is the energy actually liberated in the form of visible light and Mc^2 is the total energy of annihilation of the original star.

The fact that the total energy liberated in a supernova outburst according to our theory is at least 230 times the observed energy in the visual range of the spectrum confirms some of our conclusions published on previous occasions.[1, 2]

In Table III we list some significant energy values for the supernova in IC 4182 as functions of possible masses M of its stellar remnant.

The various energies listed have the following meaning. The total energy liberated in our supernova according to the neutron star hypothesis in the period from August, 1937 until June, 1938 is designated with E_t. Since the density of the stellar remnant presumably is not constant but increases towards its center, it is obviously

$$E_t>|E_g|, \tag{109}$$

where E_g is the gravitational energy of contraction for a sphere of mass M and constant density ρ_0 as given by (102). In the limiting case $M=M_{min}$, in which the stellar remnant has reached its final stage of contraction $\rho_0=$const. $\leqq10^{14}$ g/cm^3 it is $E_t=|E_g|$. Further, E_{max} is the energy of contraction in the case that the stellar remnant has reached its final stage $\rho_0=10^{14}$ g/cm^3. According to (3) it is

$$E_{max}=3\times10^{-3}M^{5/3}\text{ ergs}. \tag{110}$$

In the case $M=M_{min}$ it is $E_{max}=E_t$, whereas for the upper limit $M=M_l$ given by the Schwarzschild solution we have $E_{max}=0.39M_lc^2$, as given by (80). The values for Mc^2 indicate the energies which would be liberated if the mass of the star were transformed into radiation in its entirety. Finally E_{vis} is the total energy liberated in the form of light in the wave-length range $6700\text{A}>\lambda>3300\text{A}$ as we have observed it in the period from August, 1937 until June, 1938.

If, for the date of June 1, 1938, we take the mass of the stellar remnant to be $M=M(\odot)$, we may draw the following conclusions from the neutron star hypothesis. Of the maximum energy $E_{max}=9.52\times10^{52}$ ergs to be liberated ultimately through the contraction into the limiting configuration $\rho_0\leqq10^{14}$ g/cm^3 more than twenty percent, that is more than 2.16×10^{52} ergs had been released by June 1, 1938. Of this energy only 0.5 $^0/_{00}$ or about 10^{49} ergs had been radiated in form of visible light. The remaining energy $E_t'=E_t(1-5\times10^{-4})\leqq E_t$ must therefore have been liberated in the form of other energies such as invisible light, kinetic energy, heating of the stellar remnant, ionization, nuclear transformations, cosmic rays, etc. It will be of great interest to determine, by elaboration of the theory and by further observations, how E_t is to be distributed over the various forms of energy just mentioned.

The difference $E_{max}-E_t$ remains available for further evolutionary processes of the stellar remnant. Since all of the supernovae found in extragalactic systems pass out of reach of present day telescopes after the short period of a few

years at most, direct information about the final stages of supernovae can obviously be obtained only through the discovery of the remnants of supernovae in our own stellar system. In this connection the suggestion made recently by Baade[18] and Morgenroth[19] that the crab nebula had its origin in a supernova several hundred years ago, may prove to be of great importance.

0. Possible Surface Temperatures of the Stellar Remnants of Supernovae

The neutron star hypothesis of supernovae discussed in the preceding sections results in estimates of the possible radii of the stellar remnants of supernovae. The combination of these estimates with the observed visible radiation of the stellar remnants allows us to estimate lower limits of their surface temperatures. For instance the supernova in IC 4182 had an absolute magnitude $\mathfrak{M}_p \leqq -9.5$ in June, 1938. Its total radiation L at that time therefore was

$$L > 3.78 \times 10^{39} \text{ ergs/sec.} \tag{111}$$

From Table I its radius for the date mentioned must have been of the order $8 \text{ km} < r_1 < 1100 \text{ km}$. From this it follows that the effective surface temperature T_e of the stellar remnant of the supernova must have satisfied the inequality

$$4\pi r_1^2 \sigma T_e^4 \geqq L. \tag{112}$$

Substituting (93), we obtain

$$4\pi\sigma T_e^4 (\Gamma M/\kappa c^2)^2 \geqq L \tag{113}$$

or

$$T_e \geqq (L/4\pi\sigma)^{\frac{1}{4}} (\Gamma M/\kappa c^2)^{-\frac{1}{2}}, \tag{114}$$

where $\sigma = 5.72 \times 10^{-5}$ ergs/cm² sec. is the Stefan-Boltzmann radiation constant. Substitution of (90) and (111) in (114) leads to

$$T_e > 7.86 \times 10^{23}/M^{\frac{1}{2}}. \tag{115}$$

Table IV. *Lowest possible values of the effective temperature T_e of the stellar remnant of the supernova in IC 4182 in June, 1938.*

M in g	T_e in degrees abs.
2.16×10^{33}	$> 5.55 \times 10^{7}$
2×10^{33}	$> 1.76 \times 10^{7}$
3×10^{34}	$> 4.54 \times 10^{6}$
3.8×10^{34}	$> 4.03 \times 10^{6}$

[18] W. Baade, Astrophys. J. 88, 285 (1938).
[19] O. Morgenroth, Die Sterne 17, 255 (1937).

The lowest possible values of the effective temperature T_e of the stellar remnant of the supernova in IC 4182 for the stage of June, 1938 are listed in Table IV. The values for T_e thus obtained check in order of magnitude the values estimated previously from entirely different considerations,[2] an agreement which strengthens our confidence in the neutron star hypothesis.

It is obvious that with temperatures as high as given by Table IV, the emission spectra from supernovae must be of an entirely different character from those characterizing ordinary stars, whose effective surface temperatures lie in the approximate range from 1000° to 100,000°. An exhaustive study of the effects of these exceedingly high temperatures on the character of the emission of light from supernovae has been made by R. Minkowski[16] in his paper on the spectra of the two recent supernovae in IC 4182 and NGC 1003.

It should be added that in addition to the high temperatures the conditions of pressure and density on the surface of the stellar remnants of supernovae according to the neutron star hypothesis are very unusual. For instance, the acceleration of gravity on the surface of the stellar core of the supernova is, in Newton's approximation,

$$g = \Gamma M/r_1^2 = \kappa c^2/r_1 \tag{115}$$

or

$$g = \kappa^2 c^4/\Gamma M. \tag{116}$$

For $\kappa = 2 \times 10^{-27}$ and $M = M(\odot)$ we therefore have

$$g = 2.44 \times 10^{12} \text{ cm sec.}^{-2} \tag{117}$$

as compared with

$$g(\odot) = 2.7 \times 10^{4} \text{ cm sec.}^{-2} \tag{118}$$

on the surface of the sun.

Finally attention should be called to the enormous intensity of the radiation near the surface of the stellar remnants of supernovae. For the supernova IC 4182, in the stage June, 1938, the intensity of radiation at $r = r_1$ according to (111) was

$$S = L/4\pi r_1^2 > 6 \times 10^{24} \text{ ergs/cm}^2 \text{ sec.} \tag{119}$$

if the value $r_1 = 7.4 \times 10^{6}$ cm corresponding to $M = M(\odot)$ is introduced. For comparison we mention that the intensity of the radiation on

the surface of the sun is

$$S(\odot) \leqq 6.3 \times 10^{10} \text{ ergs/cm}^2 \text{ sec.} \quad (120)$$

We again refer to the recent study by Minkowski[16] where the effects of radiation intensities of the order of S on the character of the emission spectra of supernovae are discussed. The effects on the width of spectral lines as well as the fundamental importance of induced emission or "negative" absorption in the spectra of supernovae have been particularly emphasized by Minkowski.

P. Uniformity Among Supernovae

Although we possess extended series of observations for a few supernovae only, already a remarkable uniformity in the behavior of these stars begins to delineate itself. This uniformity becomes apparent through the study of the absolute magnitudes, the light curves and the spectral characteristics of supernovae.

From the available material Baade[18] has determined the mean photographic magnitude $\overline{\mathfrak{M}}_p$ of supernovae as well as the dispersion σ in $\mathfrak{M}_p$. His figures are

$$\overline{\mathfrak{M}}_p = -14.2 \quad (121)$$

and

$$\sigma \leq 1.1. \quad (122)$$

Baade also suggests that σ may be smaller yet, because the data on which the value (122) is based are partially fragmentary. Statistical studies of about 2000 films obtained with the Schmidt telescope at Palomar concerning the easily detectable occurrence of absolutely bright supernovae even at very great distances indicate in fact that the frequency of unusually bright supernovae is smaller than one should expect from a dispersion of the order of (122). The small value of the dispersion indicates a considerable uniformity in the physical behavior of supernovae. Incidentally this fact will eventually make possible the use of the apparent maximum brightness and of the shape of the light curves of supernovae as a reliable gauge of distances.

The light curves of supernovae of the same absolute brightness are remarkably similar. All of the light curves are characterized by an extended maximum of about two weeks followed by an initial decline in brightness of about 0.1 magnitude per day for the first month after the maximum. This decline coupled with a subsequent leveling off of the light curves according to Baade is definite enough to make possible the determination of the brightness at maximum, even if only relatively fragmentary light curves are available.

The evolution of the spectra of supernovae with respect to the progressive redshift of the blue bands also seems to be uniform enough to make possible the determination of the date of maximum brightness, as Minkowski's analysis[16] has revealed.

Furthermore, as Minkowski points out, the redshift about one year after maximum brightness of all supernovae seems to reach asymptotically a value of about 150 angstroms.

From the neutron star hypothesis the uniformity just described in the behavior of supernovae becomes comprehensible. If in particular, as has recently been suggested,[20] the possible masses of the neutron cores of supernovae can be limited *a priori* to a range considerably narrower than that given by (100), the uniformity of supernovae may find a very satisfactory explanation. In view of these possibilities it appears at any rate desirable to spare no efforts in the continuation of the search program for supernovae. It will be particularly important to discover supernovae at very early stages of their outbursts in the hope of securing spectra which can be interpreted in all of their essential details.

We recapitulate that the study of supernovae, in addition to the enrichment of our knowledge of the observable characteristics of this interesting class of stars, may eventually prove to be of considerable interest for the following theoretical reasons.

(a) The elucidation of the causes for supernova outbursts will mean the understanding of the generation of stellar energy for at least one class of stars. So far the problem of the generation of energy cannot be considered as having been solved satisfactorily for any class of stars.

(b) If the neutron star hypothesis of the origin of supernovae can be proved, it will be possible to subject the general theory of relativity to tests which according to the considerations presented in this paper deal with effects which in order of magnitude are large compared with the

[20] G. M. Volkoff, Phys. Rev. **55**, 421A (1939).

tests so far available. The general theory of relativity, although profound and exceedingly satisfactory in its epistemological aspects, has so far practically not lent itself to any very obvious and generally impressive applications. This unfortunate discrepancy between the formal beauty of the general theory of relativity and the meagerness of its practical applications makes it particularly desirable to search for phenomena which cannot be understood without the help of the general theory of relativity.

(c) The possibility that cosmic rays originate in supernovae constitutes an added incentive for the continued pursuit of investigations of these extraordinary stars.

My thanks are due to Professor R. C. Tolman, in discussions with whom many of the results given in the first part of this paper were derived, and to Drs. Baade and Minkowski of the Mt. Wilson Observatory in collaboration with whom all of the observational data on supernovae were critically examined on numerous occasions.

APRIL 15, 1939 PHYSICAL REVIEW VOLUME 55

Resistance, Emissivities and Melting Point of Tantalum

L. MALTER AND D. B. LANGMUIR
Research and Engineering Department, RCA Manufacturing Company, Harrison, New Jersey

(Received February 20, 1939)

The relation between true and brightness temperatures of tantalum was determined by means of pyrometric observations on the inside and outside of a long thin-walled tube heated electrically. The brightness temperature as a function of heating current was determined by sighting on an electrically heated wire of known diameter. Potential leads of fine tungsten wire welded to tantalum filaments permitted a determination of the electrical properties as a function of temperature. The electrical properties at the melting point yielded a value for this quantity from the extrapolated temperature *vs.* electrical property relations. Value for the melting point determined is 3269°K.

INTRODUCTION

MEASUREMENTS of the resistance, power radiation and brightness temperature of tantalum between room temperature and the melting point are described here. The work was undertaken when certain inconsistencies with previously published data for tantalum were discovered during experiments on the rate of evaporation of this metal which are described in an accompanying paper.[1] The relationship between resistivity and power radiation which was observed during this work is not the same as that published by Worthing.[2, 3] For a given value of power radiation the specific resistance reported here is about four percent lower than Worthing's value. It seems probable that the constitution of tantalum as made today may differ from that of the product available at the time of the earlier work. Forsythe and Watson[4] have reported an apparent change in the properties of tungsten which took place in an unexplained manner between 1922 and 1934. The earlier values of resistivity were the higher. Tantalum apparently has undergone a similar type of change.

Brightness temperature observations were also made, and are in good agreement with the earlier work both of Worthing and of Utterback and Sanderman.[5] An apparent disagreement between the latter two experimenters is due to a discrepancy between the data as shown in the table and in one of the graphs of Worthing's paper.

Except where noted the tantalum used in these experiments was regular stock from the Fansteel Metallurgical Corporation.

[1] D. B. Langmuir and L. Malter, Phys. Rev. **55**, 748 (1939).
[2] A. G. Worthing, Phys. Rev. **28**, 190 (1926).
[3] A. G. Worthing, Phys. Rev. **28**, 174 (1926).
[4] W. E. Forsythe and E. M. Watson, J. Opt. Soc. Am. **24**, 114 (1934).
[5] C. L. Utterback and L. A. Sanderman, Phys. Rev. **39**, 1008 (1932).

Procedure of Measurement

1. Brightness *vs.* true temperature

A blackbody with tantalum walls was made by rolling a narrow strip of two-mil tantalum sheet into the form of a cylinder 0.065 inch in diameter and four inches long. The edges were overlapped but not welded because it had previously been found that welding resulted in a non-uniform temperature distribution around the circumference. A number of 0.010-inch holes were drilled at various points in the middle portion of the cylinder. These served as blackbodies for the determination of the true temperature. The current leads were made of tantalum strip so proportioned as to minimize lead loss. The temperature of the cylinder was uniform over about 80 percent of its length at the lowest temperature studied.

This structure was sealed into a glass bulb and baked on exhaust at 450°C for 30 minutes. The tantalum was then outgassed by heating to 1750°K for 14 hours and by subsequent flashing for several 20-second intervals at 2600°K. After this processing, the cylinder was removed from the bulb and sealed into a new clean bulb. The purpose of this procedure was to minimize light absorption which would be due to tantalum evaporated during outgassing. The new assembly was baked at 450°C for 14 hours and the tantalum outgassed at 1750°C for five hours with seven, interposed, 30-second flashes at 2400°C. A barium getter was exploded before seal-off.

True temperatures and brightness temperatures were determined by sighting first with a telescopic, disappearing filament type of pyrometer obliquely into one of the holes, and then upon the external surrounding region. Current through the pyrometer filament was measured with a Leeds & Northrup Type K potentiometer and standard resistance. The pyrometer was calibrated by means of a wide-strip, tungsten filament lamp which was subsequently calibrated by the National Bureau of Standards for the effective wave-length $\lambda_e = 0.652\mu$. At temperatures in excess of 1800°K, an amber filter was used to reduce the observed brightness; the actual true and brightness temperatures were obtained from the values with the filter interposed by means of the relation

$$1/T_A - 1/T_0 = \lambda_e \log \tau / 0.434 C_2 = A, \quad (1)$$

where T_A is the brightness temperature, degrees Kelvin; T_0 is the apparent brightness temperature, viewed through amber filter, degrees Kelvin; C_2 is the Wien-Planck constant, 1.433 cm·degrees; and τ is the transmission of the filter at the wave-length of observation. This equation follows directly from Wien's law of radiation. The filter constant A was determined by the National Bureau of Standards for the combined amber and red eyepiece filters. The effective wave-length of the red filter was $\lambda_e = 0.674\mu$. All calibrations were corrected so as to apply to $\lambda_e = 0.665\mu$.

Several runs were taken with ascending and descending temperature values at two of the holes, but with no variation in results greater than the experimental error. The observed temperature values T_1 were corrected for the effects of wall and film absorption in accordance with the relation

$$1/T_1 - 1/T_0 = -0.0076 \times 10^{-3} \text{ degree}^{-1}.$$

The indicated value was determined experimentally after cracking open the bulb. Since this correction applies to both true and brightness temperatures, it has practically no effect on the relation between the two temperatures.

2. True temperature *vs.* heating current

A 0.012-inch tantalum wire 10 centimeters long was mounted in a bulb provided with a flat Pyrex window. Means were provided for interposing a nickel disk between the filament and window to prevent evaporation from the former depositing on the latter during exhaust. The assembly was baked on exhaust at 450°C for one hour and the filament outgassed by heating to 2300°K for 30 minutes and then to 2465°K for four minutes. A getter was exploded and the tube sealed off.

No appreciable lead loss occurred at the midpoint of this filament throughout the range studied. Brightness temperature observations as a function of current were taken in the manner indicated in the preceding section. The current values were measured with the Leeds & Northrup potentiometer. Readings were taken at suc-

cessively higher temperatures. When temperatures in excess of 2000°K were attained, the current was dropped to the value corresponding to 1750°K for a check reading before progressing to the next higher value. In this way any change due to evaporation could be detected. Actually, each reading up to the highest temperature (2952°K) used was made with sufficient rapidity so that no effect due to evaporation was observed.

A section of the filament was then weighed, and its diameter computed, by using the value 16.60 g/cm³ for the density. From these results the brightness temperature of a tantalum wire as a function of the current traversing it can be obtained.

3. Electrical properties

Eight tubes were built with tantalum filaments ranging from 0.005 inch to 0.020 inch in diameter. Each was provided with 0.002-inch tungsten potential leads. One of these tubes was built with 0.005-inch tantalum wire made in Austria about 1932. These tubes were baked for one hour at 450°C, and then the filament was outgassed for one hour at 2300°K followed by short flashes at about five minutes at 2465°K and one minute at 2630°K. Four of these tubes contained 0.012-inch diameter filaments which had previously

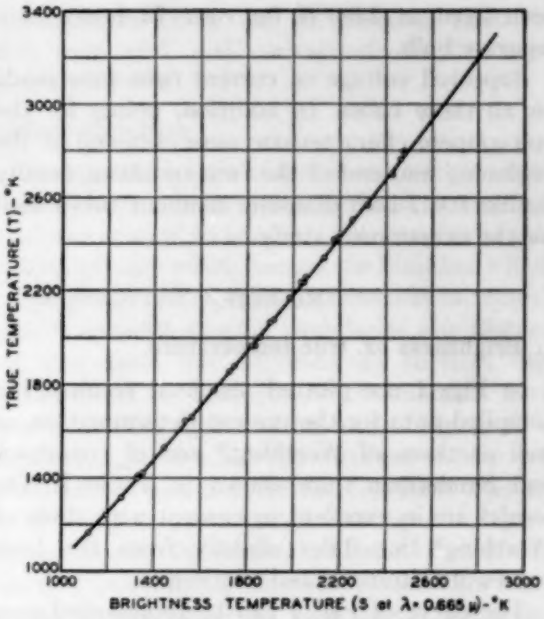


FIG. 1. True temperature T *vs.* the brightness temperature S for the effective wave-length $\lambda = 0.665\mu$.

TABLE I. *Data on resistance, emissivity, and thermal expansion of tantalum. T is the true temperature; ρ is the resistivity corrected for thermal expansion; W is the total radiation intensity not corrected for thermal expansion and e_t is the power emissivity corrected for thermal expansion; M and L are the results of Malter and Langmuir; W are those of Worthing, and U and S those of Utterback and Sanderman. The spectral emissivity at 0.665μ is indicated by $e_{0.665\mu}$. The thermal expansion given by Worthing is designated by l/l_0.*

T °K	ρ OHM-CM	W WATTS/CM²	e_t	A'	V'	$V'A'^{\frac{1}{2}}$	BRIGHTNESS TEMPERATURE M AND L	W	U AND S	$e_{0.665\mu}$	l/l_0
1000	44.1	0.793	0.136	211	0.0118	0.0702	967	966		0.481	1.0047
1100	47.3	1.23	.144	254	.0152	.0963	1060			.476	1.0055
1200	51.0	1.84	.153	299	.0193	.1291	1152	1149		.469	1.0063
1300	54.8	2.73	.163	352	.0244	.1723	1242		1250	.462	1.0071
1400	59.0	3.95	.174	408	.0304	.2255	1332	1329	1337	.456	1.0079
1500	62.4	5.47	.184	469	.0368	.2855	1421			.449	1.0087
1600	65.8	7.36	.194	528	.0438	.3540	1508	1506	1508	.442	1.0095
1700	69.3	10.10	.205	602	.0527	.4450	1596			.437	1.0103
1800	72.5	13.28	.215	676	.0617	.5415	1682	1682	1678	.432	1.0111
1900	75.8	17.12	.223	751	.0716	.6508	1767			.426	1.0119
2000	78.9	21.63	.232	828	.0821	.7709	1852	1851	1843	.421	1.0127
2100	82.0	27.11	.240	910	.0936	.9071	1933		1926	.417	1.0135
2200	85.2	34.18	.247	1002	.1072	1.080	2018	2018		.413	1.0144
2300	88.3	42.23	.254	1095	.1212	1.250	2099			.409	1.0152
2400	91.3	51.27	.261	1189	.1357	1.437	2181	2181		.405	1.0161
2500	94.4	62.38	.269	1288	.1522	1.656	2261			.402	1.0170
2600	97.4	75.37	.276	1394	.1699	1.898	2341	2339		.400	1.0179
2700	100.2	89.89	.282	1502	.1880	2.153	2421			.397	1.0188
2800	102.9	105.5	.288	1606	.2064	2.417	2499	2495		.394	1.0197
2900	105.6	123.0	.293	1715	.2257	2.699	2575			.391	1.0206
3000	108.7	144.4	.298	1830	.2479	3.032	2652	2647		.388	1.0216
3100	111.4	167.4	.302	1948	.2700	3.367	2727			.386	1.0225
3200	113.9	194.2	.306	2075	.2940	3.749	2803			.384	1.0235
3269	115.5	214.5	.309	2164	.3110	4.025	2855			.383	1.0242

been aged at 2300°K for over 72 hours in a separate bulb.

Repeated voltage *vs.* current runs were made on all these tubes. In addition, points for the volt-ampere characteristic were obtained at the beginning and end of the runs on about twenty similar 0.012-inch diameter filament tubes used for the evaporation study.

Results

1. Brightness *vs.* true temperature

In Fig. 1 are plotted the final results. The compiled data for the present determination, as well as those of Worthing,[3] and of Utterback and Sanderman,[4] are shown in Table I. The results are in excellent agreement with those of Worthing[2] but differ slightly from the later values of Utterback and Sanderman.

The curve of Fig. 1 can be represented over the range of experimental points by

$$T=0.9919S+37.14\times10^{-6}S^2+5.74\times10^{-9}S^3 \quad (2)$$

where S is the brightness temperature. The percentage error at any point is less than 0.5 percent.

This equation has been used to extrapolate the temperature scale curve to higher values for a determination of the melting point.

The spectral emissivity $e_{0.665\mu}$ as a function of T was determined from the curve of Fig. 1 by means of the relation

$$1/T-1/S=\lambda_e \log e_{0.665\mu}/0.434C_2.$$

The values were plotted and a smooth curve drawn through them. The values obtained from the smooth curve are given in Table I. The probable error of emissivity values is 0.01.

2. True temperature *vs.* heating current

The average results obtained in the manner described above are given in Table I, together with some values of Worthing. The heating current, in this case, is expressed in terms of A' where $A'=A/d_0^{\frac{1}{2}}$, $A=$the heating current, and $d_0=$the filament diameter at room temperature. While the probable error in any series of measurements on any single tube is about 0.2 percent, a spread of as much as 1.0 percent occurs between different samples of tantalum.

The spread in these values as well as the deviations from Worthing's values are very likely due to differences in the purity of the tantalum. This will be discussed in greater detail below.

3. Electrical properties

From the measured values of voltage and current, the resistivity and power radiation per cm² of tantalum were determined throughout the temperature range. The determined values were corrected for thermal expansion in accordance with the expansion data of Worthing.[4] The average values are included in Table I. In this case, too, the spread between different tubes is about 1.0 percent.

It was found that if the filament was heated in a partial pressure of air, the difference of potential across the potential leads was increased for a given value of heating current. Further heat treatment in vacuum did not restore the original values.

This fact indicates that Worthing's values which deviate in the same direction (but to a considerably greater extent) may have been obtained with less pure tantalum. That such was the case appears reasonable since the tantalum he used was made prior to 1914 and very likely inadequately "purged" during manufacture, because of the newness of the art.

The concordance of the values for the Austrian tantalum made in 1932 with those for the Fansteel product further indicates the relative purity of the present day product. Further information bears out this view. Thus, for example, the authors' value for resistivity at 300°K of 13.85×10^{-6} ohm·cm, which is lower than Worthing's value of 14.6×10^{-6} ohm·cm, is evidence of the greater purity of present day stock.

The resistivity between 0°C and 100°C can be expressed by

$$\rho=(12.56+0.048T°\text{C})\times10^{-6} \text{ ohm·cm}.$$

4. Total emissivity

The total emissivity e_t was determined from the data of Table I by means of the Stefan-Boltzmann law. These values are included in Fig. 1.

In addition, in Table I are included values of V' and $V'A'^{\frac{1}{4}}$. These, together with values of A', are of use in the design of filaments of circular cross section for particular applications, and are defined as follows:

$$V' = Vd_0^{\frac{1}{2}}/l_0, \quad V'A'^{\frac{1}{4}} = VA^{\frac{1}{4}}/l_0,$$

where V is the measured difference of potential, d_0 is the filament diameter at room temperature, and l_0 is the distance between the potential leads at room temperature.

It is of interest to note that while the use of V' or A' as determined from Worthing's data would lead to a disagreement in temperature values, the use of Worthing's values for power radiation or $V'A'^{\frac{1}{4}}$ lead to temperature values only slightly different from those of the authors. These results indicate that the presence of small quantities of impurities affect the resistivity, but not the radiation from tantalum at a given temperature.

It should be noted that the quantities A', V', $V'A'^{\frac{1}{4}}$, and W as given in Table I are not corrected for the expansion of tantalum, but are given in terms of the dimensions at 25°C. In order to permit of their correction in cases desired, the values of l/l_0 as determined by Worthing[3] are included in the table. However, the total emissivity e_t has been computed on the basis of corrections for expansion.

Melting point

The melting point of tantalum was determined by measuring the current at which a filament of known diameter burned out. Eight samples of wire of varying history were tried, and the currents at burn-out were the same within two parts in one thousand. No difference was found due to the position in which the filament loop hung. When mounted with the loop up, the filament crumpled and collapsed along several centimeters of its length. This fact combined with the reproducible values of current indicate that the determination is accurate to 0.2 percent with respect to current.

The value of A' at the melting point is 2164 amp./cm$^{\frac{1}{4}}$, and corresponds to the temperature 3269°K.

Surface structure

The appearance of the tantalum surfaces when viewed with a microscope differed somewhat from one sample to another. The outside of the hollow cylinder which formed the blackbody had a high polish and showed no scratches or roughness. A network of grain boundaries was visible, but this could not be discerned through the pyrometer. The condition of the filament surfaces depended upon the amount of aging at a high temperature which the wire had undergone. Longitudinal scratches initially present due to the drawing process smoothed out upon aging so that ultimately the surface consisted of polished, rounded, individual grains whose diameters were approximately one-tenth that of the filament. The die marks and grain boundaries both were visible under a microscope during the brightness-temperature observations which were made at a stage midway in the aging process.

When such a surface was viewed through the pyrometer, details of the surface could be seen. Ability to see the details indicated that the emissivity varies from point to point over the surface. The average emissivity, however, does not seem to be materially affected by the progressive changes in surface structure, since the volt-ampere characteristics of the filaments prove to be unaffected by aging. All aging processes were carried out using alternating current. Direct-current heating produces radical changes in the form of the surface as has been shown by Johnson.[6]

The writers are indebted to Dr. H. J. Miller for supplying a sample of Austrian-made wire, to T. B. Perkins for assistance in setting up the pyrometric apparatus, and to V. K. Zworykin and B. J. Thompson for their interest and counsel.

[6] R. P. Johnson, Phys. Rev. **54**, 459 (1938).

The Rate of Evaporation of Tantalum

D. B. LANGMUIR AND L. MALTER
Research and Engineering Department, RCA Manufacturing Company, Harrison, New Jersey

(Received March 2, 1939)

The rate of evaporation of tantalum was determined by measuring the change of resistance and the change of weight of uniform filaments. Temperatures were held constant throughout each run by adjustments of voltage V and current A so that $VA^{\frac{1}{2}}$ remained approximately unchanged. The rate of evaporation can be expressed by $\log_{10}M=7.86-39{,}310/T$, where M is the rate of evaporation in grams per cm² per sec. and T is the temperature on the Kelvin scale.

THE rate of evaporation of tantalum as a function of temperature, which appears not to have been reported previously in the literature,[1] was measured in the experiments described below. The amount of evaporation was measured by observing both the loss of weight and the increase of resistance of wires heated electrically in vacuum. The temperature scale published in an accompanying paper by the authors was used.[2]

EXPERIMENTAL METHOD

Tantalum filaments (manufactured by the Fansteel Company) having a diameter of 0.030 cm and a length of 15 cm were mounted in the shape of hairpin loops in evacuated bulbs and heated with a regulated power supply for varying lengths of time depending upon the temperature. All the wire used was aged in vacuum for times varying from 4 to 100 hours at about 2200°K before being placed in the final experimental bulbs. Tungsten wires 0.005 cm in diameter welded to the tantalum 2.5 cm from each end served as potential leads, so that measurements of voltage drop were not affected by cooling at the ends.

When temperatures were so high that the complete evaporation run lasted only a few hours or less, direct current was used for power, and values of voltage and current were followed with potentiometer measurements. For longer runs the use of direct current was impracticable because of the difficulty of obtaining a constant source of sufficient power, and so regulated alternating current was employed. With direct current, measurements of voltage V and current A were made at several stages of each evaporation. The temperature of the tube while heated by alternating current was determined by observing the current flowing in a photoelectric cell placed near the tube, and measuring the direct heating current which gave the same light output. As the ripple in light output caused by the alternating-current heating was of the order of one percent, the error due to temperature fluctuations caused by a.c. is negligible.

The mean temperature which prevailed during an evaporation was determined from the value of $VA^{\frac{1}{2}}$ which most nearly corresponded to conditions during operation. All measured values of V and A which corresponded to a.c. operating temperatures were plotted and successive points joined by a line. A straight line with slope such that $VA^{\frac{1}{2}}=$constant (i.e., an isothermal) was then drawn so that the area between it and the operating line was the same on one side as on the other. The maximum variation in temperature from the mean during any run lay between 5 and 10 degrees. This corresponds to a change in evaporation rate sufficiently small that averaging can introduce no serious error.

After about ten percent of the mass of a filament had evaporated, the tube was opened and the mass per unit length of the center portion of the wire compared with that of a sample of the batch of aged wire from which the tube was made. The change in mass was also computed from the change in resistance. This was observed from the $V-A$ plot for the operating tempera-

[1] The vapor pressures of about 30 elements are discussed in a review by C. Zwikker, Physica **8**, 241 (1928). Dr. Zwikker has informed us that one of the curves there marked "Ta" actually refers to thallium.

[2] L. Malter and D. B. Langmuir, Phys. Rev. **55**, 743 (1939).

ture. The criterion for constant temperature at beginning and end of run was that $VA^{\frac{1}{2}}$ be the same.

Results

The change in diameter of a filament as determined by weight was on the average 2.5 percent greater than the value obtained by purely electrical measurements. Differences as large as 7 percent were observed in individual cases. As these discrepancies are definitely larger than the experimental error for either type of measurement, some change of the tantalum during the evaporation runs seems to be indicated. This change amounts to about 0.2 percent of the value of the resistivity. The variation of resistivity among supposedly identical samples of tantalum was found during work on the tantalum temperature scale[1] to be of this same order of magnitude.

The rate of evaporation as a function of temperature as determined from weight measurements is plotted in Fig. 1. The curve can be represented by the equation:

$$\log_{10} M = 7.86 - 39{,}310/T,$$

in which M is the rate of evaporation in grams/cm² sec. and T is the temperature on the Kelvin scale. Values of evaporation rate and vapor pressure for 200-degree temperature intervals are given in Table I. The vapor pressure was calculated from the evaporation rate using the equation:[2]

$$P = 1700 M T^{\frac{1}{2}},$$

in which it is assumed that there is no lessening of the evaporation rate caused by reflection of atoms at the surface.

TABLE I. *Values of the evaporation rate and vapor pressure of tantalum at various temperatures.*

T °K	M g/cm² sec.	P dynes/cm²
2000	1.63×10^{-12}	1.27×10^{-7}
2200	9.78×10^{-11}	8.01×10^{-6}
2400	3.04×10^{-9}	2.58×10^{-4}
2600	5.54×10^{-8}	4.90×10^{-3}
2800	6.61×10^{-7}	6.07×10^{-2}
3000	5.79×10^{-6}	5.40×10^{-1}
3200	3.82×10^{-5}	3.77
3269 (M.P.)	6.80×10^{-5}	6.75

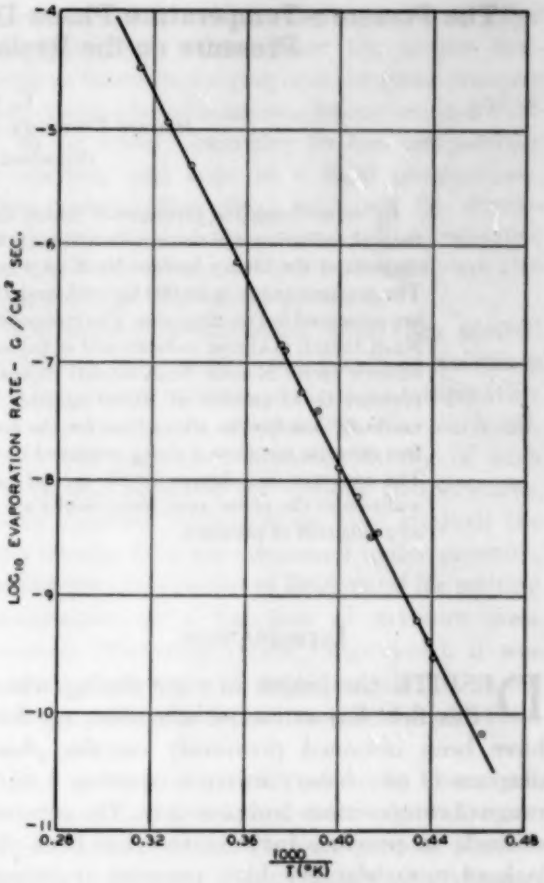


FIG. 1. Rate of evaporation of tantalum as a function of temperature.

The authors express their appreciation of the advice and interest of Mr. B. J. Thompson and Dr. V. K. Zworykin.

[1] H. S. Taylor, *Treatise on Physical Chemistry* (D. Van Nostrand Company, second edition, 1931), Vol. I, p. 245.

The Pressure-Temperature Phase Diagram of Na-K Alloys and the Effect of Pressure on the Resistance of the Liquid Phase

C. H. KEAN
*Harvard University, Cambridge, Massachusetts**

(Received March 6, 1939)

By determining the pressures at which discontinuities occur in resistance isotherms of mixtures of sodium and potassium in various proportions, it has been possible to map out the phase diagram of the binary system Na-K as a function of pressure, temperature and composition. The pressure range is 10,000 kg/cm², and the temperature range from 0° to 150°. The results are presented in two diagrams. The composition of the eutectic in the subsystem (solid Na-solid Na_2K-liquid) is almost independent of pressure over the range, whereas the composition of the eutectic (solid K-solid Na_2K-liquid) moves from about 67 atomic percent K at atmospheric pressure to 59 percent at 10,000 kg/cm². The rise of melting temperature with pressure is markedly less for the alloys than for the pure metals: the rise of melting temperature for the first eutectic mentioned above produced by 10,000 kg/cm² is 42°, against 68° for pure sodium. The corresponding figures for the second eutectic are 42° against 105° for pure potassium. In addition to the phase data, some results are given for the relative resistance of the liquid alloys as a function of pressure.

INTRODUCTION

DESPITE the length of time during which this field has attracted attention, no data have been obtained previously on the phase diagram of any binary mixture covering a wide range of compositions and pressures. The greatest obstacle to previous investigators has been the lack of a satisfactory high pressure technique which now, however, has reached a state of perfection—sufficient to permit handling of pressures up to 10,000 kg/cm², the range of this work, with considerable ease.[1]

In all, it appears that about 20 binary mixtures[2] have been investigated with special emphasis usually placed on the relation between the composition of eutectic and either of the possible independent variables, pressure or temperature. Most of the mixtures examined were organic compounds or cryohydrates; only one system of metals, Na-Hg, has been studied[3] and no variation of the composition of the eutectic mixture over the pressure interval 1–4000 kg/cm², was found. With the exception of the recent work of Deffet[2] most data were obtained with only a limited number of compositions, and hence it was necessary to complete the phase diagram by unjustified extrapolations.

METHOD

Since the usual high pressure apparatus can be subjected to only moderate temperatures, it was necessary to select a binary system which was composed of metals with conveniently situated melting points. The metals, sodium and potassium, satisfy this condition. In addition, they possess a relatively simple phase diagram at atmospheric pressure.

The alloys of different compositions were prepared by mixing weighed amounts of the metals. The original metals, which were obtained from the J. T. Baker Chemical Company, were heated in a vacuum in order to remove absorbed gases; and then run through a series of constrictions, which mechanically cleaned them of their insoluble oxides, into tubes which were sealed and weighed. The ends of these tubes were later broken off and then they were placed immediately into a third vessel which was sealed and evacuated. On heating, the metals melted and ran together and were mixed by rocking the vessel; the connection with the vacuum pump was through a heavy rubber tube. Next the alloy was run into several U-tubes which were sealed and set aside for subsequent use; care was taken

* Now employed by the Humble Oil and Refining Company, Houston, Texas.

[1] P. W. Bridgman, *The Physics of High Pressure* (Macmillan, 1931).

[2] Louis Deffet, Soc. Chem. Belgique **45**, 213 (1936); this article contains an excellent bibliography.

[3] N. A. Puschin and J. V. Grabenschtschikow, Zeits. f. physik. Chemie **118**, 276 (1925).

in this operation to have the alloys liquid and hence homogeneous. The composition was ascertained by weighing the fragments of the original containers of the metal which were recovered from the mixing vessel.

The metals prepared by the above procedure were found to have a purity satisfactory for these experiments. The purity of the original metals was tested by measuring the resistance at atmospheric pressure as a function of temperature through a temperature range covering the melting point. The criterion of purity was the abruptness of the change of resistance on melting, a criterion which is known to be very sensitive to dissolved impurities, the most important impurities for this work. For these measurements, the molten metal was run in vacuum through constrictions into a long glass tube through which four platinum wires were sealed; the resistance was measured on a potentiometer. The sodium gave an exceptionally sharp melting curve; the potassium showed some "rounding" but the melting point was 0.3° higher than that in the literature.[4]

The resistance measurements of the alloys were made with thin-walled tubes of soft glass about two millimeters in internal diameter and 9–10 cm long bent in a U-shape for compactness within the pressure cylinder. Two short side tubes of equal diameter were sealed near the top of the U and through these platinum wires 0.003 inch in diameter were sealed, which served as the potential leads to the potentiometer. If heavier platinum leads were used there was prohibitive cracking of the glass because of differential compressibility. Larger bore tubes sealed to the ends of the U-tube served as a reservoir for the highly compressible alloys. The filled and sealed U-tubes were prepared for the resistance measurements by breaking off the ends and inserting fine helical, amalgamated copper wires which served as the current leads to the potentiometer. The tubes were immersed in a clear mineral oil during this operation to prevent oxidation. The copper showed no signs of deterioration over the period used—normally several weeks.

The U-tubes were next attached to the usual three-terminal plug and inserted in the high pressure cylinder which was immediately filled with petroleum ether. After the proper connections between the plug and the potentiometer were made, the cylinder was immersed in a bath of oil or water, according to the temperature of interest, and kept at a fixed temperature. The pressure was then increased by convenient amounts and the corresponding resistance changes were noted; resistance changes were also recorded with decreasing pressure.

The resistance isotherms were then plotted from these data; that is, the curves of resistance as a function of pressure at constant temperature and constant gross composition. For each different composition a different family of such curves was obtained for various temperatures. Seven different compositions were studied; the pure metals were not measured under pressure, but the previous results of Bridgman[5] for melting temperature as a function of pressure were assumed. The temperature range which it was necessary to cover with each composition varied with the composition, and may be inferred from the diagrams.

The location of crucial points on the phase diagram may be deduced from the resistance curves. As long as a single homogeneous phase is present, resistance varies smoothly with pressure, but if pressure crosses into a region of two-phase equilibrium, a second phase will, in general, begin to separate out, and there will be a discontinuity in the slope of the resistance isotherm. The phase diagram as a function of pressure is to be determined from these points of discontinuity in slope. The points of discontinuity may be grouped and plotted in various ways, as for example pressure as a function of temperature at constant composition and the phase diagram determined by extrapolation or interpolation of suitable curves. The actual carrying out of this general scheme encountered various complications because of sluggishness in the attainment of equilibrium, arising from subcooling of the liquid or from the diffusion which has to take place after every change of pressure in a two-phase region, because the composition of the two phases is, in general, different and a function of pressure. In spite of the diffusion, which would introduce

[4] P. W. Bridgman, Phys. Rev. 3, 153 (1914).

[5] P. W. Bridgman, Phys. Rev. 27, 68 (1926).

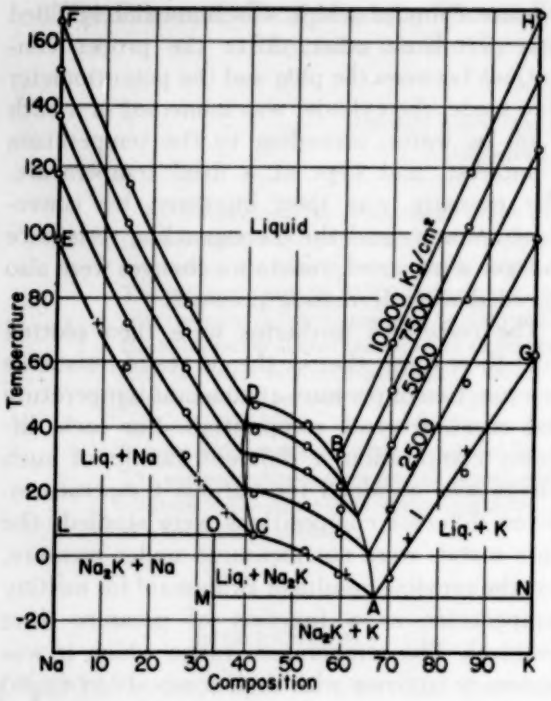


FIG. 1. Phase diagram of the binary system Na-K.

a hysteresis between readings with increasing and decreasing pressure, it was found that the rate of diffusion was sufficiently high, even when one of the phases was solid, so that usually consistent readings with increasing and decreasing pressure could be attained without too tedious delays.

The various details of supercooling, etc., obviously depend on the precise phases present, and can be further discussed when the complete results are before us.

RESULTS

The most important results of this paper are presented in Fig. 1, which is the phase diagram of the system under study extended to a pressure of 10,000 kg/cm². Here the phase diagrams at five different pressures (1, 2500, 5000, 7500 and 10,000 kg/cm²) have been projected on the composition-temperature plane. The composition is given in atomic percent potassium. The circles represent data obtained as suggested by replotting and interpolation or extrapolation. The crosses represent data at atmospheric pressure taken from a paper by E. Rinck.[6]

[6] E. Rinck, Comptes rendus, 197, 49 (1933).

Considering the diagram as a whole we find at high temperatures a region of unlimited extent in which the system is liquid. As the temperature is lowered or the pressure increased a surface consisting of three smooth intersecting surfaces is reached which represents the equilibrium between the liquid phase and a solid phase; the solid phase is either Na, K, or the intermetallic compound Na_2K. The discontinuities obtained on going through the smooth surface (liquid +Na, liquid), the sheet *CEFD* in Fig. 1, and (liquid+K, liquid), sheet *ABHG*, were reversible; i.e., there is no appreciable subcooling when solid pure Na or pure K separate from the liquid phase. In the remaining case the curve had to be extended for some distance into the region liquid+Na_2K before the solid Na_2K appeared; however, once started, the precipitation ran quite rapidly. Since the converse reaction, solid to liquid, does not superheat, these transition points were determined only from the curve obtained with decreasing pressure. For the other surfaces the transitions were determined from the curves drawn for increasing pressure.

Next, the surfaces which establish the equilibrium between three phases were obtained and consequently by the phase rule must be capable of generation by a line moving parallel to the composition axis; namely, the lines *LC* and *MN* move up along the lines *CD* and *AB* with increasing pressure, generating two surfaces. The composition corresponding to the line *OM* presumably stays constant since this represents a molecular compound. These surfaces indicate the equilibrium in one case between liquid, Na and Na_2K; in the other between liquid, Na_2K and K. On passing through these surfaces a discontinuous change of resistance is experienced, corresponding to the abrupt disappearance of one phase. The transitions in the former case were irreversible and it was generally necessary to advance the pressure several thousand kg/cm² before Na_2K began to precipitate. Since this is a case of subcooling, it is evident that the point found with decreasing pressure was the true pressure at which thermodynamic equilibrium existed. The amount of subcooling was a function of the previous history of the alloy. For instance, if the pressure on the virgin specimen was advanced in steps from one atmosphere the

formation of the third phase might occur suddenly at a pressure 1000 or 1500 kg/cm² above the true transition point. The transition would not occur at a pressure below this even though pressure and temperature might be held fixed for as much as 12 hours. However, if the transition was once allowed to run and then pressure decreased below the true transition point but not to atmospheric pressure and then increased again the transition now ran at a pressure intermediate between the true and the previous "false" transition pressure.

The curves of Fig. 2 represent the pressure-temperature traces of the developable surfaces and indicate the equilibrium between three phases on the temperature-pressure plane; the curves obviously also represent the temperature-pressure dependence along the lines *AB* and *CD* of Fig. 1. The points of the curves were averaged from the results for a wide range of compositions and were not confined solely to compositions in the neighborhood of the lines *AB* and *CD*. The points at 1 atmosphere pressure were obtained from the paper by Rinck.[6] Similar curves were drawn and studied for the intersection of planes of constant composition with the two phase equilibrium surfaces. These curves are not presented here in detail but they can be reproduced from Fig. 1. In general, both these curves for constant composition and the curves of Fig. 2 for three-phase equilibrium are like the usual melting curve; namely, they are concave to the pressure axis with the values of $(\partial t/\partial p)_s$ decreasing with increasing pressure. However, some of the curves for the two-phase surfaces appear to violate this, for at low pressure the curves seemed to be slightly convex to the pressure axis; the evidence is not conclusive. It is evident from Fig. 1 that, in general, the values of $(\partial t/\partial p)_s$ for the two-phase surfaces exceed those for the three-phase surfaces at a given pressure. In general, on both sheets *CEFD* and *ABHG*, the average value of $(\partial t/\partial p)_s$ decreases markedly as one adds the other component to the pure metal.

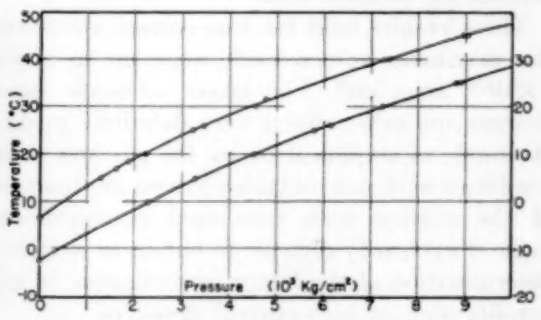


FIG. 2. Lower curve gives equilibrium between liquid, K and Na_2K, i.e., temperature as a function of pressure along the line *AB* of Fig. 1; temperature axis to right. Upper curve gives equilibrium between liquid, Na and Na_2K, i.e., temperature as a function of pressure along the line *CD* of Fig. 1; temperature axis to left.

TABLE I. *Relative resistances of four representative compositions of* Na-K *liquid alloy.*

PRESSURE KG/CM²	15.0% K 115°C	27.3% K 80.0°C	52.9% K 40.0°C	85.5% K 120°C
1	1.000	1.000	1.000	1.000
1000	0.937	0.954	0.960	0.913
2000	0.892	0.917	0.925	0.844
3000	0.855	0.887	0.896	0.788
4000	0.825	0.862	0.873	0.743
5000	0.799	0.840	0.854	0.706
6000	0.776	0.822	0.838	0.677
7000	0.756	0.806	0.825	0.655
8000	0.738	0.793	0.814	0.634
9000	0.721	0.781	0.805	0.617
10000	0.705	0.771	0.798	0.601

Apart from their use in establishing the phase diagram, it is evident that the resistance measurements of this paper have an interest in themselves. However, this interest is mostly confined to the measurements on the single liquid phase, previous experience having shown that reliable measurements of the pressure coefficients of solids even as soft as these cannot be made in glass capillaries because of the constraints exerted by the glass. Also, since the dimensions of the glass capillaries were not determined, the results can be used only to determine relative resistances. Table I gives the results of representative compositions of the liquid at the indicated temperatures. The relative resistances are "as measured" and are uncorrected for compressibility of the glass capillary. It is to be noted that these results are quite characteristic for metals and alloys. Also, it is apparent on comparison with other work that the change of resistance with pressure is much less than for the corresponding pure substances.[7] There is good evidence for the existence of a pressure minimum of resistance for sodium and the minimum for

[7] P. W. Bridgman, Proc. Am. Acad. 72, 157 (1938).

potassium is known to occur at 25,000 kg/cm²;[7] a plot of the first differences of Table I suggests that the same phenomenon occurs at least with the alloys rich in potassium and probably at pressures lower than for the pure potassium.

Finally, it may be of interest to mention that certain indirect evidence indicates that alloys of about 50 percent K must possess highest compressibilities.

The author wishes to acknowledge the assistance of Professor Bridgman who suggested the problem and took an interest in its solution.

APRIL 15, 1939 PHYSICAL REVIEW VOLUME 55

Townsend Ionization Coefficients in Cs-Ag-O Photo-Tubes Filled with Argon*

W. S. HUXFORD
Northwestern University, Evanston, Illinois
(Received August 15, 1938)

The attempt is made to account for the amplification of photo-currents at low current densities in argon-filled Cs-Ag-O tubes on the basis of Townsend's theory assuming secondary electron emission by positive ions. Special sealed-off tubes were used having parallel plates at various separations and a constant gas pressure. Values of the Townsend coefficients α and γ were determined at each value of field by simultaneous solution of the Townsend equations for two plate separations. The α/p_0 *vs.* E/p_0 plot has a maximum at $E/p_0 \leqq 700$ volts/cm·mm. The γ-values for the two types of cathodes used are sensibly constant at low fields. Measured values of the striking potentials of the glow discharge are in quantitative agreement with the positive ion mechanism assumed. The photosensitivity of the compound caesium cathodes was found to increase with increase of bombarding ion current.

STUDIES of the amplification of photoelectric currents by gases are complicated by the circumstance that several auxiliary processes may act simultaneously to augment the primary ionization by electron impact. In commercial photo-tubes having compound low work-function cathodes, it is especially difficult to determine which of the possible secondary mechanisms are operative and to measure their relative effects. Although numerous measurements have been made in tubes in which metallic cathodes are used, little information exists as to the numerical value of the Townsend coefficients in tubes having activated photo-cathodes. An attempt has been made to obtain such information, and in this paper results are reported for Cs-Ag-O tubes filled with argon which indicate: (1) that the photosensitivity of the compound cathode increases with increase in positive ion current to the cathode; (2) that in gas-amplified currents at all potentials secondary mechanisms play a significant part; (3) that the Townsend coefficient α for argon increases with increasing field to a maximum, and then decreases; and (4) that the chief auxiliary mechanism at low current densities leading to high gas amplification, and accounting quantitatively for the low values of striking potential is the release of electrons at the cathode by positive ions.

These results hold for low current densities; the maximum values used range up to about 7×10^{-8} amp./cm². For larger currents space charges and other effects very definitely modify the mode of amplification by the gas and these combined with unavoidable changes in structure of the coating with continued bombardment make it extremely difficult to obtain an accurate determination of the factors contributing to gas amplification at high current densities.

Photo-Tubes and Apparatus

In view of the difficulty of maintaining uniform emissivity from a caesium cathode over a long

* A preliminary report of this work was made at the 222nd regular meeting of the American Physical Society, Toronto, Canada, on June 24, 1938. W. S. Huxford, Phys. Rev. **54**, 313(A) (1938).

period of time while the tube is on the pumping system, several sealed-off tubes were prepared in a uniform way with the same gas pressure. Six tubes were used having parallel plates with separations of 2.0, 4.2, 6.0, 8.1, 12.3 and 16.0 mm. These tubes were made by the Research and Engineering Department of the RCA Manufacturing Company, following a special design. They were filled with pure argon at a pressure of 0.175 mm Hg at room temperature. The bulbs were tubular, about 2.5 inches in diameter, and the plates were plane square sheets of nickel and silver measuring 1 inch on a side. A circular aperture 0.5 inch in diameter was cut in the nickel sheet used as anode and the entire inside surface was covered with a nickel mesh having openings 0.4 mm square. The silver plate which served as cathode was oxidized in the usual way, coated with caesium, treated and baked in the manner employed in the preparation of commercial photo-tubes. The cathode could be illuminated by light sent from outside the tube through the aperture in the anode. The mesh transmitted approximately 50 percent of the light falling upon it.

Photosensitivity curves for all of the caesium oxide cathodes were obtained by means of a Van Cittert double monochromator with a glass dispersing system used with a tungsten filament source. The wave-length spread of the monochromator with a 0.2-mm slit width was about 8A at 4000A and increased to 100A at 6000A. The relative amounts of energy in the light beams from the monochromator were determined by means of a vacuum thermocouple. Currents were measured by two Leeds and Northrup galvanometers each having a sensitivity of the order of 5×10^{-11} ampere per mm. Considerable care was taken in checking the linearity of response of these galvanometers. Some of the currents were measured by balancing the photocurrent with a potentiometer circuit, while the galvanometer was used as a null instrument.

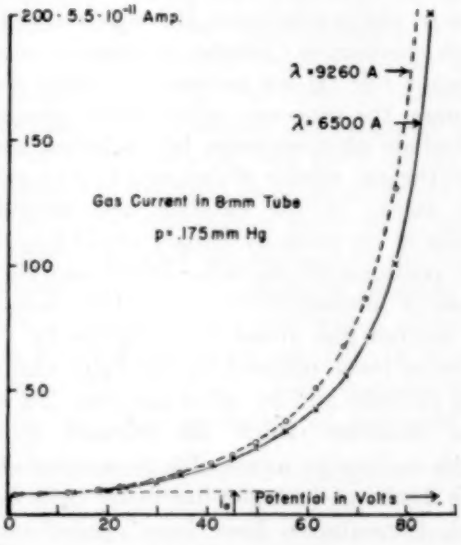


FIG. 1. Dependence of gas current upon the tube potential for two monochromatic beams of light in which the intensity of light has been adjusted to give the same value of i_0, the initial photo-current.

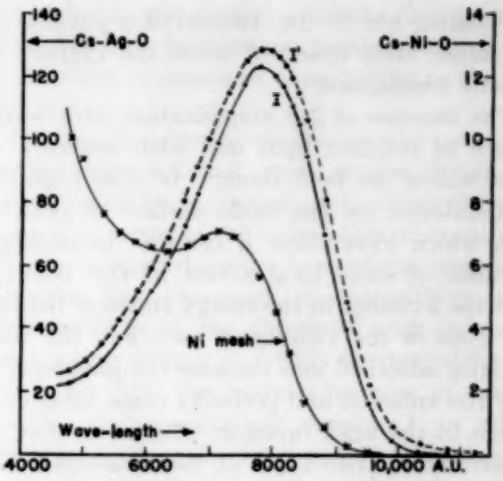


FIG. 2. Photosensitivity plots for the silver and nickel electrodes showing in arbitrary units the current per unit radiant energy as a function of wave-length. Curve I gives discharge currents in argon per unit energy in the 12-mm tube at 54 volts, and I' the currents due to the same monochromatic beams when a discharge current of 0.1 microampere due to an auxiliary white source was set up simultaneously in the tube. The other curve shows the relative sensitivity of the activated nickel electrode.

CHANGE OF PHOTOSENSITIVITY WITH ION CURRENT

When measuring the photoelectric response of the cathodes while currents were passing in the gas, it was discovered that the amplification was noticeably greater for long wave-length radiations than for shorter wave-lengths. For large emission currents this effect was somewhat masked by the general increase in amplification which occurs with increase in intensity of illumination. In the low current range this effect is not due to a change in mechanical structure of

the coating nor to the removal of a gas film as has often been observed when the cathode is heavily bombarded.

The increase of gas amplification with wave-length of incident light and with intensity of illumination are both thought to arise from the accumulation on the oxide surface of positive ions which have there a certain "mean life." Because of some local action in the coating, perhaps a change in the energy states of the free electrons in the vicinity of each ion, the temporarily adsorbed ions increase the likelihood of electron emission and probably cause local variations in the work function. This effect may be illustrated by two types of measurements. The curves of Fig. 1 show the amplification obtained when light of wave-lengths 9260A and 6500A falls on the oxide coating. At the higher potentials the amplification produced by the longer-wave-length is from 20 to 30 percent higher than for the shorter. This could be explained by a shift of the photoelectric threshold to the red.

A second type of critical experiment is illustrated in Fig. 2. The solid curve (I) shows the gas currents per unit light intensity when only resolved light is used. The dashed curve (I') shows the enhanced sensitivity produced in the cathode when a simultaneous current is set up in the gas (by an auxiliary source of white light) which is approximately seven times the maximum current obtained with the monochromator alone. The effect is reproducible and is present only during the bombardment.

Figure 3 shows in detail amplification curves in the low current range for all the tubes. Log i/i_0 is here plotted as a function of the accelerating potential, where i_0 is the saturation current obtained at potentials less than 18 volts (see Fig. 1). In all cases distinctly larger amplifications are obtained with infra-red light ($\lambda \geqq 9000$A, dashed curves) than with red light ($\lambda \geqq 6600$A, solid curves). These curves are for small current densities ($<10^{-8}$ amp./cm^2) unless otherwise indicated.

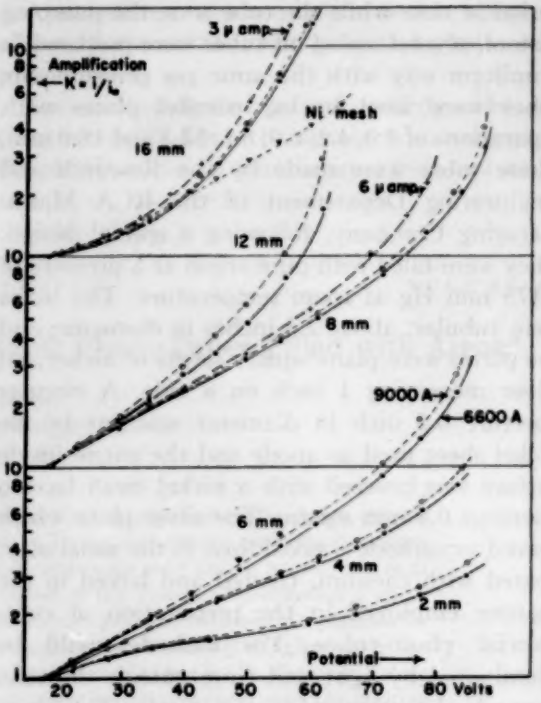


FIG. 3. These plots show the gas amplifications of low intensity photo-currents in argon-filled Cs-Ag-O tubes when the cathode is illuminated by two monochromatic beams of different color. The wave-lengths in each case are approximately those shown for the 16- and 8-mm tubes, with white light, which indicate the increase in amplification at high current densities.

DISCHARGE MECHANISMS INVOLVED

An examination of the plots of Fig. 3 on the basis of Townsend's theory shows that the amplification depends to a considerable degree upon processes other than simple ionization of the gas by electron impacts. It is well known that in such discharges cathode phenomena play a significant role. If we assume the argon to be very pure, the processes which merit consideration include photoemission by radiation generated in the gas, release of electrons by metastable argon atoms at the cathode and secondary emission of electrons by impact of positive ions.

For pressures of the order of 0.1 mm Hg the number of photoelectrons released by radiation from excited gas atoms is negligible by comparison to those released by the light admitted to the cathode and by other agencies. To what extent electrons which are released from a cathode surface by metastable atoms contribute to the discharge in argon has never been determined. Experiments have been carried out in this laboratory on the time required for the discharge in argon-caesium photo-tubes to reach a state of equilibrium after light is flashed upon

the cathode[1] which indicate that a part of the measured lag may be attributed to the relatively slow diffusion of metastable atoms. It is well known that these atoms effectively ionize minute quantities of vapor or gaseous contamination when the energy of the metastable state exceeds that of the ionization potential of the impurity. In the absence of such contamination the only contribution to the discharge would be the release of electrons from the cathode surface. A number of investigators believe that the time lag in Townsend currents in rare gases is to be ascribed chiefly to the action of metastable atoms. An exception to this interpretation is the recent measurements by Skellett[2] of the change in amplification with frequency of the incident light pulse in the case of a spherical photo-tube filled with argon. He finds no evidence that metastable atoms contribute to ionization processes in the gas, and accounts for the observed dynamic response of the photo-currents in terms of the transit time of argon ions which release secondary electrons upon arrival at the cathode.

Kingdon and Thomson[3] among others have demonstrated that secondary emission of electrons by positive ion bombardment of Cs-Ag-O cathodes exists, and that it plays a major role in gas amplification in argon-filled photo-tubes. This mechanism was first proposed by J. J. Thomson and by J. S. Townsend to account for the abnormally large current amplifications observed when an alkali cathode was employed in the discharge tube. That the effect is particularly important in compound cathodes is shown by the work of Guntherschulze and Betz[4] on the glow discharge. For example, these investigators found that in a glow discharge in pure argon, cathode fall 1000 volts, each ion released on the average 1.88 electrons from a cathode of oxidized magnesium; while for a pure magnesium surface the value was 0.37 electrons per ion. In a recent report on the effect of cathode material on the second Townsend coefficient in pure and contaminated N_2 gas Bowls[5] concludes that emission of electrons by positive ion bombardment is one of the chief mechanisms contributing to the discharge.

During the course of the present study it became apparent that many of the effects observed could be explained in a straightforward manner in terms of the interaction of argon ions with the activated surface layer of the compound cathode. The enhanced photosensitivity and observed shift of the threshold with increased bombarding current is readily accounted for on this basis. Also the marked increase of amplification at large current densities,[6] an effect which we have so far been unable to explain in terms of a simple increase of positive ion space charge in the neighborhood of the cathode, may be thought of as due to an increase in charge density of argon ions temporarily adsorbed on the cathode surface and resulting in an enhance-

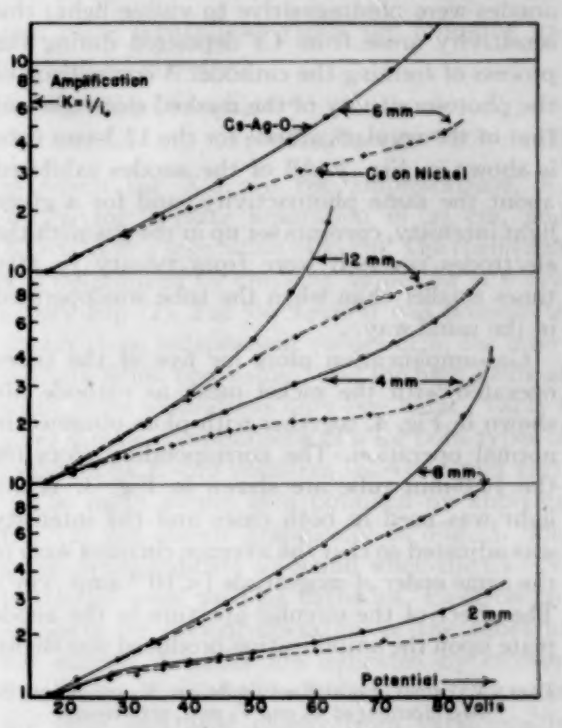


FIG. 4. Amplification plots at low current densities in each tube for the caesium oxide cathode (solid curves) and for the activated nickel mesh (dashed curves). The corresponding plots for the 16-mm tube are shown in Fig. 3.

[1] See W. S. Huxford and R. W. Engstrom, Rev. Sci. Inst. **8**, 385–390 (1937) for details of the circuit used in measuring time lags in argon photo-tubes.

[2] A. M. Skellett, J. App. Phys. **9**, 631–634 (1938).

[3] K. H. Kingdon and H. E. Thomson, Physics **1**, 343–351 (1931).

[4] A. Guntherschulze and H. Betz, Zeits. f. Physik **108**, 780–785 (1938).

[5] W. E. Bowls, Phys. Rev. **53**, 293 (1938).

[6] The apparent increase in amplification at current densities of 3μ amp./cm^2 and at 6μ amp./cm^2 is shown in Fig. 3.

ment of the photoelectric efficiency of the emitting layers. All of these effects are small at low current densities, and it seemed worth while to carry out measurements of the Townsend coefficients under the simplest possible conditions at very low light intensities in order to determine, if possible, whether the hypothesis of release of electrons from the cathode by positive ions provides an auxiliary mechanism which adequately accounts for the amplifications observed in the absence of other complicating phenomena.

Photo-currents with Nickel Mesh Cathodes

Studies of secondary emission of electrons by positive ions show that this process is highly sensitive to changes in structure of the emitting surface. Fortunately it was possible in the present work to vary the cathode surface simply by reversing the role of the two electrodes. Preliminary tests had shown that the nickel mesh anodes were photosensitive to visible light; this sensitivity arose from Cs deposited during the process of forming the cathode. A comparison of the photosensitivity of the meshed electrode and that of the regular cathode for the 12.3-mm tube is shown in Fig. 2. All of the anodes exhibited about the same photoactivity, and for a given light intensity, currents set up in the gas with the electrodes reversed were from twenty to fifty times smaller than when the tube was operated in the usual way.

Gas-amplification plots for five of the tubes operated with the nickel mesh as cathode are shown in Fig. 4, together with plots obtained in normal operation. The corresponding plots for the 16.0-mm tube are shown in Fig. 3. White light was used in both cases and the intensity was adjusted so that the average currents were of the same order of magnitude ($<10^{-8}$ amp./cm^2). The effect of the circular aperture in the anode plate upon the amplification produced was slight.

Table I. *Values of α at $E=50$ volts/cm. K_1 and K_2 are the amplifications at V_1 and V_2 volts, respectively.*

Cathode	x_1	x_2	V_1	V_2	K_1	K_2	α	α_{av}	α/p_0	E/p_0
Cs-Ag-O	0.89	0.47	61.5	40.5	16.0	2.35	0.84			
"	0.47	0.26	40.5	30.0	2.35	1.55	1.00			
"	0.26	0.08	30.0	21.0	1.55	1.14	0.88	0.91	5.7	312.5
Cs-Ni	0.89	0.47	61.5	40.5	6.0	2.20	1.16			
"	0.47	0.26	40.5	30.0	2.20	1.50	1.12			
"	0.26	0.08	30.0	21.0	1.50	1.13	1.00	1.09	6.7	312.5

Sensibly the same results were obtained when light was focused on an area of the mesh backed by the solid nickel plate as when it was focused on the mesh over the half-inch hole. Some differences were found when the whole mesh was diffusely illuminated and when the light was concentrated on a small area.

The branching of the two curves for each plate separation furnishes a most interesting and instructive comparison of the effect of reversing the role of the two electrodes. Obviously the efficiency of ionization of the gas by electron impact is the same in any tube at a given potential, so that the striking differences in amplification which are observed must depend upon an auxiliary mechanism associated with the cathode (or possibly the anode). A crucial test of the hypothesis proposed to account for the amplification in these tubes is furnished by applying the equation of Townsend and determining the value of the ionization coefficient for electrons at each value of the electric field for both normal and reversed polarity of the electrodes. The value of this coefficient should be sensibly constant for both cases at each field value, while the secondary Townsend coefficients should exhibit systematic differences for the two types of electrodes. Such calculations have been carried out and are presented in the sections which follow.

Calculation of α-Values

The equation of Townsend for gas-amplified photo-currents in which secondary electrons are produced by positive ions at the cathode is given by

$$i=i_0\frac{e^{\alpha x}}{1+\gamma-\gamma e^{\alpha x}}, \qquad (1)$$

in which α represents the number of pairs of ions produced by each electron per cm of path, γ is the average number of electrons released at the cathode by each positive ion reaching it, and $x=d-\delta=d-V'/E$ is the effective plate separation for a parallel plate arrangement. The current i_0 is the photo-current generated by radiation at the cathode and is measured at potentials so low that ionization of the gas by electrons does not occur (see Fig. 1); and δ is the average distance an electron moves from the cathode in order to

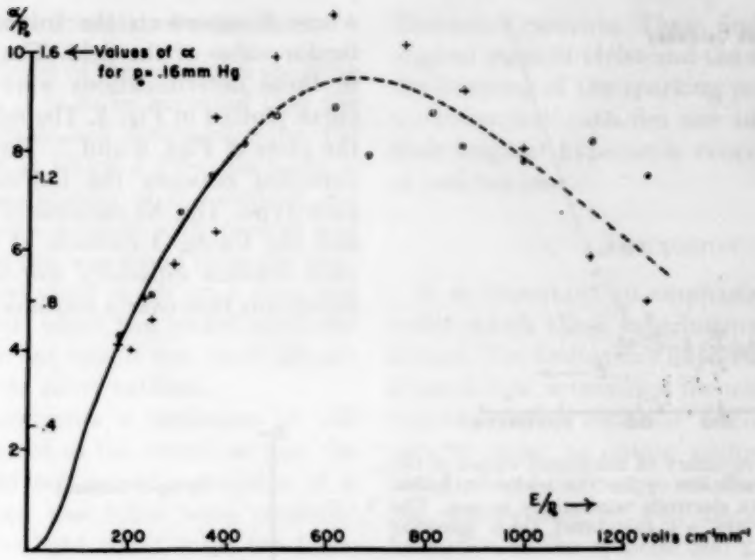


FIG. 5. Values of α/p_0 for the nickel cathodes are shown by small circles, those for the Cs-Ag-O cathodes by crosses. The solid curve represents α/p_0 for pure argon as obtained by Kruithof and Penning.

acquire enough energy to ionize a molecule of the gas upon impact. An approximate value is given by $\delta = V'/E$, where V' is the effective ionizing potential of the gas in volts,[7] and E is the electric field in volts per cm.

The quantities α and γ may be calculated from Eq. (1) for a particular value of field by a simultaneous solution involving two plate separations, x_1 and x_2. This method would be expected to be particularly effective in cases, such as the present, where γ is large even at low fields. For the six tubes used in this study, however, the results are not particularly impressive, because γ does not have the same value for presumably identical cathodes at a given field intensity. There is the additional complication that the photosensitivity of the activated silver cathodes changes with the magnitude of the bombarding ion current. One important result of the present calculations is that the values of the Townsend coefficients at high fields ($E/p > 500$) are the first to be published for argon in discharges of this type.

If we can assume that at a given field strength γ is constant for cathodes prepared in the same way, the simultaneous solution of Eq. (1) for two plate separations yields the following expression involving α:

$$K_1(K_2-1)+(K_1-K_2)e^{\alpha x_1} = K_2(K_1-1)e^{\alpha(x_1-x_2)}, \quad (2)$$

in which the K's are the amplifications, $K_1 = i_1/i_0$ for a plate separation of x_1 cm, and $K_2 = i_2/i_0$ at a separation of x_2 cm. The coefficient α is determined by choosing values for it which satisfy Eq. (2). For low fields it was possible to obtain three independent values of α, for intermediate fields two values, and for high fields at small plate separations only one value. Table I gives sample computations of α for both types of cathodes at $E = 50$ volts/cm ($p_0 = 0.16$ mm Hg).

The mean values of α obtained by applying Eq. (2) are shown graphically in Fig. 5. The small circles indicate α's found when the Cs−Ni electrodes were used as cathodes. Electron coefficients for normal operation of the tubes at low current densities are indicated by crosses. The fluctuations of these values from the mean are large and arise chiefly from variations in structure of the activated silver oxide. There is a definite correlation between these fluctuations and the activation treatment given the different cathodes. The surface structure in the 12.3-mm tube, for example, is distinctly different from that in the 8.1-mm tube.

[7] A value of $V' = 17$ volts was used in the present calculations and is an estimate based on the best agreement between data on the various tubes. A. A. Kruithof and F. M. Penning (Physica **3**, 524 (1936)) used $V' = 15.7$ volts in their measurements of α in pure argon.

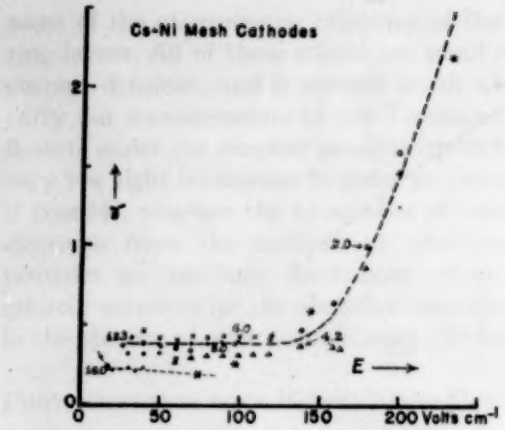


FIG. 6. Graphical summary of calculated values of the second Townsend coefficient γ for the nickel cathodes. The numbers indicate electrode separations in mm. The points marked (•) are γ's calculated from sparking potentials.

The solid curve in Fig. 5 gives values for argon obtained by Kruithof and Penning.[8] These investigators give no data beyond $E/p_0=429$ volts/cm·mm. The dashed curve represents the trend of α/p_0 as a function of E/p_0 for higher fields as determined in the present investigation. It is evident that computations of α in discharges with the nickel cathode show more uniformity than those obtained with normal polarity of the electrodes. Although α values beyond $E/p_0=1000$ volts/cm·mm are unreliable, since these depend solely upon one computation involving the 4.2-mm and 2.0-mm tubes, it is believed that α/p_0 for argon reaches a maximum at about $E/p_0=700$ and then decreases at higher fields. This conclusion requires confirmation under conditions in which the pressure of the gas can be varied. The marked constancy of the γ values in this region, as shown in the following section, tends to support the reality of this variation of α with field.

VALUES OF THE SECOND TOWNSEND COEFFICIENT, γ

The coefficient γ was calculated at various fields for each of the twelve electrodes used as cathode. Solving Eq. (1) for γ we have

$$\gamma=\frac{K-e^{\alpha x}}{K(e^{\alpha x}-1)}, \qquad (3)$$

where K represents the amplification at a particular value of the field. The values of α used in these determinations were taken from the curve plotted in Fig. 5. The results are shown in the plots of Figs. 6 and 7. There is considerable variation between the individual cathodes of each type. The Ni cathode of the 16-mm tube and the Cs-Ag-O cathode of the 8.1-mm tube each exhibits especially low values of γ. It is significant that over a considerable range of field

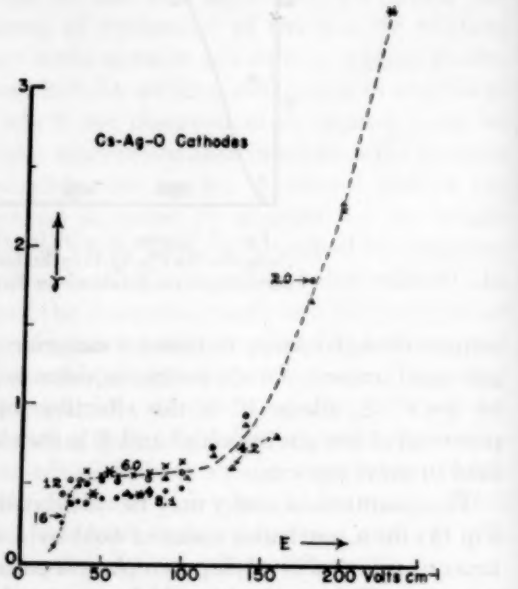


FIG. 7. Graphical summary of calculated values of the second Townsend coefficient γ for the Cs-Ag-O cathodes. The numbers indicate electrode separations in mm. The points marked (•) are γ's calculated from sparking potentials.

strengths the values of γ for most of the cathodes are sensibly constant. For the nickel cathodes this range includes the region in the neighborhood of $E=110$ volts/cm at which α reaches its maximum value. The present results cannot be taken as indicating accurately the form of the upward trend of γ at high fields, since the values in this region are least reliable. It is to be expected, however, that at fields sufficiently high the efficiency of the ions in releasing secondary electrons should increase rather rapidly as the ion is able to gain appreciable amounts of energy from the field in one mean free path. This sudden increase in γ is verified by the following data obtained from measurements of the potentials at which a glow discharge is set up in the gas.

[8] A. A. Kruithof and F. M. Penning, Physica **3**, 529 (1936); **4**, 447 (1937).

γ-Values from Striking Potentials

In the neighborhood of the minimum sparking potential of a gas it is to be expected that the nature of the cathode will noticeably alter the value of potential at which a glow discharge is initiated. In measurements of the striking potentials in the tubes studied with normal and reversed electrodes this was found to be the case, for the potential required to set up a glow was considerably higher when the nickel electrode was the cathode; in one tube it was nearly double that required for the silver cathode.

In these measurements a resistance of 460 megohms was inserted in the circuit so that the glow current could not exceed a fraction of a microampere. When the tubes were carefully shielded so that no light could enter the bulb, the error in determining the glow-initiating potential was about 2 volts. The values of the potentials so determined are shown in Table II for both types of cathodes.

According to the Townsend theory of sparking, if Eq. (1) represents correctly the manner in which the gas currents grow, the condition for disruptive discharge is the increasing of i to an infinite value. In this case we have

$$1+\gamma-\gamma e^{\alpha x}=0, \quad \text{or} \quad \gamma=(e^{\alpha x}-1)^{-1}. \qquad (4)$$

Values of the positive ion coefficient calculated by means of this equation from the α corresponding to the value of field at which striking of the glow occurred, are shown in the columns marked γ_s of Table II. For the 4.2- and 2.0-mm tubes γ's cannot be computed because the values of α are not known at very high fields.

The values of γ_s given in Table II are indicated by an asterisk (*) in the plots of Figs. 6 and 7, and are in good agreement with the ion coefficients obtained from measurements of the Townsend currents. These findings confirm the original ideas of Holst and Osterhuis[9] concerning the lowering of the sparking potential when low work-function cathodes are used, and support their original hypothesis concerning the action of positive ions.

TABLE II. *The γ-values from striking potentials.*

Electrode separation	Nickel V_s	Cathode γ_s	Silver V_s	Cathode γ_s
16.0 mm	118.0 v	0.185	72.4 v	0.43
12.3	120.3	0.26	63.5	0.58
8.1	117.0	0.60	89.5	0.58
6.0	140.0	(2.20)*	92.5	1.20
4.2	129.0		95.0	(3.50)*
2.0	147.6		97.0	

* These estimates were obtained by extrapolating the α/p_0 curve of Fig. 5.

Discussion

It is important to emphasize the conditions under which these experiments have been conducted. The limitations imposed included the use of small light intensities, the use of plane parallel electrodes with emission restricted to a small area in order to obtain uniform fields and to avoid effects of charges on the bulb walls, and the employment of well tested procedures in formation of the cathode and purification of the filling gas. It is believed that the precautions taken were justified in view of the excellent reproducibility of photosensitivity and gas amplification which was obtained throughout the course of the investigation. No attempt has been made to secure a complete analysis of gas amplification at high current densities in argon photo-tubes as employed in technical work.

The agreement of the values of α/p_0 at low fields with those of Kruithof and Penning indicate that the argon used was relatively free from contamination by foreign gases or vapors. The large variation in values of the electron ionization coefficient obtained at higher fields, however, indicates that not all of the complicating effects previously mentioned have been eliminated. Some of the fluctuations undoubtedly arise from unavoidable variations in cathode activation and gas pressure in the six tubes used. Differences in γ-values for the same type of cathode surface in different tubes indicate considerable fluctuation in the secondary emission efficiencies of the individual cathodes. Part of the lack of agreement in α-values obtained with normal and reversed polarity of the electrodes may be caused by the ionization of traces of Cs vapor or by release of electrons from the cathodes by metastable argon atoms. That positive ions play an important role in the amplification process, however, seems to be well established by the results

[9] G. Holst and E. Oosterhuis, Physica **1**, 78 (1921); Phil. Mag. **46**, 1117 (1923).

obtained in this attempt to apply Townsend's theory to a type of discharge which is complicated by many factors.

It is believed that the enhancement of cathode photosensitivity by ion bombardment indicates a lowering of the cathode work-function by the joint action of adsorbed ions and applied field. This "cathode" effect becomes increasingly important at high current densities where distortion and hysteresis occur in normal operation of commercial tubes and where time lag effects are very large. The study of gas currents in these tubes at higher current densities is being continued with particular reference to the dynamical characteristics of the discharge.

In conclusion the writer wishes to acknowledge the help of the Research and Engineering Department of the RCA Manufacturing Company, Inc., Harrison, New Jersey and also support for the project from the Graduate School of Northwestern University in the form of a grant. Mr. Dwight Wennersten gave valuable assistance in taking data for the photosensitivity plots and Dr. A. M. Glover of the RCA Manufacturing Company was responsible for the preparation of the tubes.

APRIL 15, 1939 PHYSICAL REVIEW VOLUME 55

The Propagation of Electric Currents in Terminated Lines

Solutions of the Telegraphic Equation*

R. H. KENT
Aberdeen Proving Ground, Maryland

(Received February 17, 1939)

By a transformation of Heaviside's general solution, a special solution for the current having the value $e^{-t}f(t)$ at $x=0$ is obtained. From this solution, an expression is obtained for the potential difference across the line at $x=0$ in terms of $f(t)$ and its successive integrals. From this and the terminal conditions a differential equation of negatively infinite order for $f(t)$ is obtained. This equation can be solved to any given degree of approximation by a method of integral operators. An illustrative problem is solved. An integral solution of the telegraphic equation and a corresponding expression for the potential difference at $x=0$ are given. An alternative method to Heaviside's and Carson's is suggested for obtaining the general solution for the terminated line.

THERE have been discovered several methods of computing the current in a terminated line when an electromotive force is applied in one of the terminal networks. Of these the principal ones are Heaviside's[1] operational methods which have been given a rigorous deductive foundation by Bromwich[2] and Carson.[3] Heaviside's solutions hold formally only for a unit electromotive force applied at time $t=0$ but Carson[3] has shown that the response to an arbitrary e.m.f. may be computed from Heaviside's results by the application of Duhamel's superposition theorem. Methods similar to those employed by Bromwich in his deduction of Heaviside's solutions have also been applied to the solution of circuit problems by Wagner[4] and Fry.[5]

The method about to be described is of the classical type, in that a special solution of the telegraphic equation is obtained. This solution involves a function $f(t)$ which is conveniently determined from the boundary conditions with an arbitrary e.m.f. by operational methods.

* Published by permission of the Chief of Ordnance, U. S. A.

[1] O. Heaviside, *Electromagnetic Theory* (1893), Vol. 1, 2 and 3. See also V. Bush, *Operational Circuit Analysis with an Appendix by N. Wiener* (1929); and E. J. Berg, *Heaviside's Operational Calculus* (1936).

[2] T. J. I'A. Bromwich, Proc. Lond. Math. Soc. (2) **15**, 401–448 (1916). See also H. Jeffries, *Operational Methods in Mathematical Physics* (1931).

[3] J. R. Carson, *Electric Circuit Theory and Operational Calculus* (1926).

[4] K. W. Wagner, Archiv. f. Electrotechnik, **4**, 159 (1916).

[5] T. C. Fry, Phys. Rev. **14**, 115 (1919).

Incidentally, it should be noted that the telegraphic equation and the equation for the damped motion of a string have the same form.[6]

Special Solution of the Telegraphic Equation

Symbols are adopted as follows: a = velocity of propagation of electric waves, c = leakage conductance, h = length of line, i = current in the line, k = capacity of line per unit length, l = inductance of line per unit length, r = resistance of line per unit length, t = time, V = potential difference between the wires, x = distance along the line from the origin.

The current, i, is propagated along the line in accordance with the partial differential equation

$$\frac{\partial^2 i}{\partial t^2}+\left(\frac{r}{l}+\frac{c}{k}\right)\frac{\partial i}{\partial t}+\frac{ci}{lk}=\frac{1}{lk}\frac{\partial^2 i}{\partial x^2}$$

with a similar equation for V.

We shall for convenience assume that the insulation is high so that $c=0$ although the solution with $c\neq 0$ is not much more difficult.

We thus obtain for i,

$$\frac{\partial^2 i}{\partial t^2}+\frac{r}{l}\frac{\partial i}{\partial t}=\frac{1}{kl}\frac{\partial^2 i}{\partial x^2}. \tag{1}$$

If the substitution $i=e^{-bt}u$ is made, where $b=r/2l$, Eq. (1) may be reduced to

$$\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2}+b^2u, \tag{2}$$

in which $a^2=1/kl$.

Heaviside's[7] solution of (2) is

$$u=f(x+at)+f(x-at)$$
$$+\frac{1}{a}\int_{-at}^{at} f(x+\lambda)\frac{\partial}{\partial t}I_0[b(t^2-\lambda^2/a^2)^{\frac{1}{2}}]d\lambda$$
$$+\frac{1}{a}\int_{-at}^{at} g(x+\lambda)I_0[b(t^2-\lambda^2/a^2)^{\frac{1}{2}}]d\lambda, \tag{3}$$

for the initial conditions,

$$\left.\begin{aligned} u&=2f(x)\\ \partial u/\partial t&=2g(x)\end{aligned}\right\}\text{at } t=0.$$

In (3) $I_0[\ ]$ is the Bessel's function of zero order with imaginary argument.

Substitutions are now made;

$$t'=x,\quad x'=t. \tag{4}$$

With these, (2) becomes

$$\frac{\partial^2 u}{\partial t'^2}=\frac{1}{a^2}\frac{\partial^2 u}{\partial x'^2}-\frac{b^2}{a^2}u=\frac{1}{a^2}\frac{\partial^2 u}{\partial x'^2}+\left(\frac{ib}{a}\right)^2 u, \tag{5}$$

where $i=(-1)^{\frac{1}{2}}$.

From (3), it appears that the solution of (5) is

$$u=f(x'+t')+f(x'-t')$$
$$+a\int_{-t'/a}^{t'/a} f(x'+\lambda)\frac{\partial}{\partial t'}I_0\left[\frac{ib}{a}(t'^2-\lambda^2a^2)^{\frac{1}{2}}\right]d\lambda$$
$$+a\int_{-t'/a}^{t'/a} g(x'+\lambda)I_0\left[\frac{ib}{a}(t'^2-\lambda^2a^2)^{\frac{1}{2}}\right]d\lambda,$$

subject to the initial conditions $u=2f(x')$, $\partial u/\partial t'=2g(x')$ at $t'=0$.

If $J_0(x)$ is Bessel's function of a real argument, $I_0(x)=J_0(ix)$. It follows that $I_0(ix)=J_0(-x)=J_0(x)$, since J_0 is an even function.

By reversing the substitutions (4), we find that

$$u=f(t+x)+f(t-x)$$
$$+a\int_{-x/a}^{x/a} f(t+\lambda)\frac{\partial}{\partial x}J_0\left[\frac{b}{a}(x^2-\lambda^2a^2)^{\frac{1}{2}}\right]d\lambda$$
$$+a\int_{-x/a}^{x/a} g(t+\lambda)J_0\left[\frac{b}{a}(x^2-\lambda^2a^2)^{\frac{1}{2}}\right]d\lambda, \tag{6}$$

subject to the initial conditions, $u=2f(t)$ and $\partial u/\partial x=2g(t)$ at $x=0$.

We now take $a=b=1$, which is equivalent to taking a/b as the unit of length and $1/b$ as the unit of time. We also assume $g(t)=0$. The solution (6) becomes

$$u=f(t+x)+f(t-x)$$
$$+\int_{-x}^{x} f(t+\lambda)\frac{\partial}{\partial x}J_0[(x^2-\lambda^2)^{\frac{1}{2}}]d\lambda. \tag{6a}$$

[6] See A. Kneser, *Die Integralgleichungen und ihre Anwendungen in der Mathematischen Physik* (1911), p. 61.

[7] Electrical Papers (1892), Vol. II, p. 478. See also A. G. Webster's *Electricity and Magnetism* (1897), p. 545. This solution was given by Heaviside without explanation of his method of deduction. The solution was subsequently derived by H. Poincaré, Comptes rendus **117**, 1027 (1893); E. Picard, *ibid.* **118**, 16 (1894); and J. Boussinesq, *ibid.* **118**, 162 (1894).

The solution (6a) is now broken into two parts designated by $u(t-x, x)$ and $u(t+x, -x)$ where

$$u(t-x, x)$$

$$= f(t-x) - \int_{-t}^{-x} f(t+\lambda) \frac{\partial}{\partial x} J_0[(x^2-\lambda^2)^{\frac{1}{2}}] d\lambda$$

$$= f(t-x) - \int_0^{t-x} f(z) \frac{\partial}{\partial x} J_0[(x^2-[z-t]^2)^{\frac{1}{2}}] dz, \quad (7)$$

$$u(t+x, x)$$

$$= f(t+x) + \int_{-t}^{x} f(t+\lambda) \frac{\partial}{\partial x} J_0[(x^2-\lambda^2)^{\frac{1}{2}}] d\lambda$$

$$= f(t+x) + \int_0^{t+x} f(z) \frac{\partial}{\partial x} J_0[(x^2-[z-t]^2)^{\frac{1}{2}}] dz. \quad (8)$$

That (7) and (8) are solutions of (2) may be shown by direct substitution.

The series expansion of $J_0[(x^2-[z-t]^2)^{\frac{1}{2}}]$ is now substituted in (7) and differentiated. The integral occurring in (7) is then expanded by successive integration by parts. In writing out this expansion it is convenient to employ integral operators, I^i, defined by the following:

$$If(y) = \int_0^y f(z) dz, \quad I^2 f(y) = \int_0^y \int_0^z f(z) dz^2,$$

$$I^i f(y) = \int_0^y [I^{i-1} f(z)] dz.$$

In the integrations the powers of t disappear with the result that

$$u(t-x, x) = \left\{ 1 + \frac{x}{2} I + \frac{1}{8}(x^2 I^2 + x I^3) + \frac{1}{48}(x^3 I^3 + 3x^2 I^4 + 3x I^5) + \frac{1}{384}(x^4 I^4 + 6x^3 I^5 + 15x^2 I^6 + 15x I^7) + \frac{1}{3840}(x^5 I^5 + 10x^4 I^6 + 45x^3 I^7 + 105x^2 I^8 + 105x I^9) + \frac{1}{46080}(x^6 I^6 + 15x^5 I^7 + 105x^4 I^8 + 420x^3 I^9 + 945x^2 I^{10} + 945x I^{11}) + \frac{1}{645120}(x^7 I^7 + 21x^6 I^8 + 210x^5 I^9 + 1260x^4 I^{10} + 4725x^3 I^{11} + 10395x^2 I^{12} + 10395x I^{13}) + \cdots \right\} f(t-x). \quad (7a)$$

The operators in the bracket of (7a) operate on $f(t-x)$ only. If the function of I and x given by the bracket of (7a) is designated by $\psi(I, x)$, (7a) may be written

$$u(t-x, x) = \psi(I, x) f(t-x). \quad (7b)$$

The corresponding expression for $u(t+x, -x)$ is found to be

$$u(t+x, -x) = \psi(I, -x) f(t+x). \quad (8b)$$

It is seen that $u(t-x, x)$ consists of waves traveling to the right (toward $+x$) and $u(t+x, -x)$ of waves traveling toward the left. The wave trains proceeding to the right and left, respectively, are independent of each other.

If $u = \psi(I, x) f(t-x)$ is a solution of (2) then $i = e^{-t}\psi(I, x) f(t-x)$ is a solution of (1). If $f(t) = 0$ for $t \leqq 0$, then this solution satisfies the following boundary conditions:

At $x=0$, $i = e^{-t} f(t)$; for $t=0$, and $x \geqq 0$, $i=0$.

To distinguish it from the reflected currents we add subscript 0's. In other words,

$$i_0(t, x) = e^{-t}\psi(I, x) f_0(t-x). \quad (9)$$

Operational Solution for Terminated Line

We assume a line in equilibrium for $t=0$, with terminals at $x=0$ connected to the net work A as shown in Fig. 1. In a branch of the network A an electromotive force $E(t)$ is applied at a time $t>0$. It is desired to determine $f_0(t)$ in terms of $E(t)$.

Let the potential difference across the line caused by the current $i_0(t, x)$ be $V_0(t, x)$. If $V_0(t, 0)$ can be evaluated* in terms of $f_0(t)$, then these expressions can be inserted in the equations obtained from Kirchhoff's laws which become equations for $f_0(t)$.

At time t, the head of the waves $f(t-x)$, etc. has not reached the point $x=t$. For $x \leqq t$, by

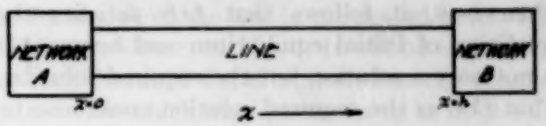


Fig. 1. Line and networks.

hypothesis $V_0=i_0=0$. Hence $V_0(t, t)=0$. However,

$$V_0(t, t)=V_0(t, 0)+\int_0^t \frac{\partial V_0}{\partial x}dx,$$

from which

$$V_0(t, 0)=-\int_0^t \frac{\partial V_0}{\partial x}dx=-\int_0^t \left(\frac{l\partial i_0}{\partial t}+ri_0\right)dx. \quad (10)$$

In (10), i_0 from (9) is now inserted and the integration carried out by parts. The result is,

$$V_0(t, 0)=e^{-t}\left\{\begin{aligned}&(r-l)\left[I+\frac{1}{2}I^2+\frac{1\cdot 3}{2\cdot 4}I^5+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}I^7+\cdots\right]f_0(t)\\&+l\left[1+\frac{1}{2}I^2+\frac{1\cdot 3}{2\cdot 4}I^4+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}I^6+\cdots\right]f_0(t)\end{aligned}\right\}$$

$$=e^{-t}l\left[1+I+\frac{1}{2}(I^2+I^3)+\frac{1\cdot 3}{2\cdot 4}(I^4+I^5)+\cdots\right]f_0(t), \quad (11)$$

since $b=r/2l=1$. Since $[1+I+\frac{1}{2}(I^2+I^3)+\cdots]$ is the series expansion of $(1+I)(1-I^2)^{-\frac{1}{2}}$, (11) may be written

$$V_0(t, 0)=e^{-t}l[(1+I)(1-I^2)^{-\frac{1}{2}}]f_0(t). \quad (11a)$$

Since $i_0(t, 0)=e^{-t}f_0(t)$, the impedance of the infinite line is

$$l[(1+I)(1-I^2)^{-\frac{1}{2}}]=l[(1+I)^{\frac{1}{2}}(1-I)^{-\frac{1}{2}}]. \quad (12)$$

To determine $f_0(t)$, Kirchhoff's laws are applied to the terminal network A. In this application, i_0 is represented by $e^{-t}f_0(t)$ and derivatives of $f_0(t)$ by $I^{-1}f_0(t)$, etc. Integrals of $e^{-t}f_0(t)$ are expanded by parts and integrals of $f_0(t)$ are represented by $If_0(t)$, etc. If this procedure is carried out and the impedance of the infinite line is represented by the series expansion of (12), an equation of the following form will be obtained for $f_0(t)$.

$$\left(\sum_{i=-n}^{\infty} a_iI^i\right)f_0(t)=\left(\sum_{i=-m}^{l} c_iI^i\right)E(t)e^t, \quad (13)$$

in which m, n and l are positive integers. Both sides of (13) are divided by the polynomial operator

$$\sum_{i=-n}^{\infty} a_iI^i$$

and the quotient,

$$\sum_{i=-m} c_iI^i \Big/ \sum_{i=-n}^{\infty} a_iI^i,$$

is expanded in ascending powers of I. (We assume this expansion is possible.) If this expansion be represented by $Q(I)$ in which $Q(I)$ is an infinite series in the ascending powers of I, a solution of (13) is

$$f_0(t)=Q(I)E(t)e^t. \quad (14)$$

This may be shown by substituting (14) in (13). Since $E(t)e^t$ and all its integrals are zero

* The ratio $V_0(t, 0)/i_0(t)$ is the impedance of the infinite line. We are therefore attempting to get an expression for the impedance in the general case.

when $t=0$, it follows that $f_0(t)$ satisfies the condition of initial equilibrium and hence (14) is not only a solution but the required solution. That (14) is the required solution could also be shown by the methods employed by Bromwich or by Carson to establish the correctness of Heaviside's results.

When the waves $f(t-x)$, etc. reach the network B at $x=h$ (see Fig. 1) they are in general reflected at least partially. We proceed to determine the current, i_1, which is the reflection of i_0.

For i_1, we take (see Eq. (8b))

$$i_1(t, x)=e^{-t}\psi(I, -x+2h)f_1(t+x-2h). \quad (15)$$

The calculation of $f_1(t)$ in terms of $f_0(t)$ may be more easily performed if new coordinates s, z, defined by $s=t-h$, $z=x-h$, are introduced. In terms of these we have

$$i_0(t, h)=e^{-(h+s)}F_0(s); \quad i_1(t, h)=e^{-(h+s)}F_1(s).$$

$F_1(s)$ is determined in terms of $F_0(s)$ by a procedure analogous to that employed in the determination of $f_0(t)$ in terms of $E(t)$. The impedance of the infinite line for currents traveling to the right is $l[(1+I)^{\frac{1}{2}}(1-I)^{-\frac{1}{2}}]$ as before while the impedance for the waves traveling to the left, $i_1(t, h)$, is $-l[(1+I)^{\frac{1}{2}}(1-I)^{-\frac{1}{2}}]$. Hence we have,

$$V_0(t, h)=e^{-(h+s)}l[(1+I)^{\frac{1}{2}}(1-I)^{-\frac{1}{2}}]F_0(s),$$
$$V_1(t, h)=-e^{-(h+s)}l[(1+I)^{\frac{1}{2}}(1-I)^{-\frac{1}{2}}]F_1(s).$$

By virtue of these relations, $F_1(s)$ may be determined in terms of $F_0(s)$ by the application of Kirchhoff's laws to the terminal network at $x=h$. The result is expressed in the form

$$F_1(s)=X_A(I)F_0(s), \quad (16)$$

in which $X_A(I)$ is a polynomial of infinite degree in I. It may be called the operational reflection coefficient for the terminal network at $x=h$. However,

$$e^{-(h+s)}F_0(s)=e^{-t}\psi(I, h)f_0(t-h),$$
$$e^{-(h+s)}F_1(s)=e^{-t}\psi(I, h)f_1(t-h).$$

From these and (16) it follows that

$$f_1(t)=X_A(I)f_0(t)=[X_A(I)][Q(I)]E(t)e^t.$$

When i_1 reaches $x=0$, it is in turn reflected; the current from the second reflection is

$$i_2(t, x)=e^{-t}\psi(I, x+2h)f_2(t-x-2h). \quad (17)$$

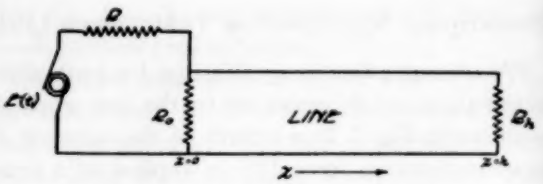


FIG. 2. Network for illustrative example.

The function $f_2(t)$ is given by

$$f_2(t)=[X_0(I)]f_1(t)=[X_0(I)][X_A(I)]f_0(t)$$
$$=[X_0(I)][X_A(I)][Q(I)]E(t)e^t,$$

in which $X_0(I)$ is the operational reflection coefficient for the network A. The operator polynomials may be multiplied together before operating on $E(t)e^t$.

By proceeding in this way, it is found that

$$f_k(t)=X_0^{\frac{1}{2}[k-\frac{1}{2}+(-1)^k\frac{1}{2}]}X_A^{\frac{1}{2}[k+\frac{1}{2}-(-1)^k\frac{1}{2}]}[Q(I)]E(t)e^t.$$

By virtue of this result, the total current, i, including all reflections is given by

$$i(x, t)=e^{-t}\sum_{k=0}^{\infty}X_0^{\frac{1}{2}[k-\frac{1}{2}+(-1)^k\frac{1}{2}]}X_A^{\frac{1}{2}[k+\frac{1}{2}-(-1)^k\frac{1}{2}]}$$
$$\times[Q(I)]\psi(I, z_k)E(t-z_k)e^{t-z_k}$$

in which

$$z_k=(-1)^k\{x+(-1)^k[k+\tfrac{1}{2}-(-1)^k]\tfrac{1}{2}h\}. \quad (18)$$

The method of obtaining X_0, X_A and $Q(I)$ has been explained. The function ψ is defined by Eqs. (7a) and (7b).

ILLUSTRATIVE EXAMPLE

Equation (18) is applied to the calculation of the current in the resistance R_A of the network shown in Fig. 2, in which the lines are in equilibrium when the electromotive force $E(t)$ is applied at $t>0$.

By applying Kirchhoff's laws to the network at $x=0$, the following equation is obtained:

$$RR_0f_0(t)e^{-t}+(R+R_0)V_0(t, 0)=R_0E(t).$$

If the expression (11a) for $V(t, 0)$ is inserted, the equation for $f_0(t)$ is

$$\left[1+\frac{(R_0+R)l}{RR_0}(1+I)^{\frac{1}{2}}(1-I)^{-\frac{1}{2}}\right]f_0(t)=\frac{E(t)e^t}{R}.$$

If we now take $(R_0+R)/R_0R=1/l$, the solution is

$$f_0(t)=\frac{1}{1+(1+I)^{\frac{1}{2}}(1-I)^{-\frac{1}{2}}}\frac{E(t)e^t}{R}$$

$$=\frac{1}{I}[(I-1)+(1-I^2)^{\frac{1}{2}}]\frac{E(t)e^t}{2R}.$$

Of course by the operator function on the right we mean its series expansion. Accordingly,

$$f_0(t)=\left[1-\left(\frac{1}{2}I+\frac{1}{2\cdot4}I^3+\frac{1\cdot3}{2\cdot4\cdot6}I^5+\frac{1\cdot3\cdot5}{2\cdot4\cdot6\cdot8}I^7+\cdots\right)\right]\frac{E(t)e^t}{2R}$$

$$=Q(I)E(t)e^t=[1-P(I)]E(t)e^t/2R, \quad (19)$$

if $P(I)$ represents the operator series in the parenthesis. In a similar manner it can be shown that if $R_k=l$

$$X_k(I)=X_0(I)=P(I).$$

The current in the terminal resistance with $R_k=l$ is therefore

$$i(t,h)=e^{-t}\left\{\sum_{k=0}^{\infty}[\psi(I,y_k)]\times[P^{2k}-P^{2k+2}]\frac{E(t-y_k)e^{t-y_k}}{2R}\right\}, \quad (20)$$

where $y_k=(2k+1)h$ and $P(I)$ is given by (19).

It is interesting to obtain the limiting form of (20) as h approaches zero. Under these conditions $\psi=1$ and $y_k=0$, the reflections being all simultaneous. We obtain

$$i(t,h)=e^{-t}[1-\lim P_{k=\infty}^{2k+2}]\frac{E(t)e^t}{2R}.$$

But $$e^{-t}[\lim P_{k=\infty}^{2k+2}]E(t)e^t=0,$$

if $E(t)$ is finite.

Hence, $$\underset{h=0}{i(t,h)}=\frac{e^{-t}E(t)e^t}{2R}=\frac{E(t)}{2R}.$$

This is in agreement with the steady-state condition.

Calculations of $i(t,h)$ have been made with $E(t)=1$ for $t>0$ and $R=\frac{1}{2}$ for various values of h, the length of the line in the stated units. The results are shown on Fig. 3. It will be noted that in each instance $i(t,h)$ approaches the steady-state value which may be shown to be $1/(1+h)$.

INTEGRAL EXPRESSIONS FOR $i_0(t,x)$ AND $V_0(t,x)$

By virtue of Eq. (7) we may write in place of (9)

$$i_0(t,x)=e^{-t}\left\{f_0(t-x)-\int_{-t}^{-x}f_0(t+\lambda)\frac{\partial}{\partial x}\times J_0[(x^2-\lambda^2)^{\frac{1}{2}}]d\lambda\right\}. \quad (21)$$

From (10) it is seen that $V_0(t,0)$ may be evaluated if

$$\int_0^x i_0dx \quad \text{and} \quad \int_0^x\frac{\partial i_0}{\partial t}dx$$

are known. In view of the fact that

$$\frac{\partial}{\partial x}\left\{\int_{-t}^{-x}f_0(t+\lambda)J_0[(x^2-\lambda^2)^{\frac{1}{2}}]d\lambda\right\}=-f_0(t-x)+\int_{-t}^{-x}f_0(t+\lambda)\frac{\partial}{\partial x}J_0[(x^2-\lambda^2)^{\frac{1}{2}}]d\lambda,$$

it follows that

$$\int_0^x i_0(t,x)dx=-e^{-t}\int_{-t}^{-x}f_0(t+\lambda)J_0[(x^2-\lambda^2)^{\frac{1}{2}}]d\lambda+e^{-t}\int_{-t}^{0}f_0(t+\lambda)I_0(\lambda)d\lambda.$$

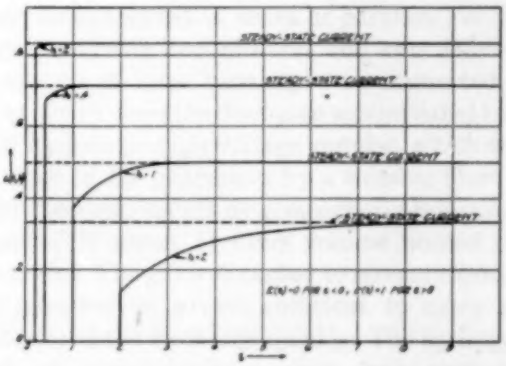


FIG. 3. The current in R_k vs. t for various lengths of line.

From this relation and a corresponding one for

$$\int_0^x \frac{\partial i_0}{\partial t} dx,$$

it may be shown, if $b=r/2l=1$, that

$$V(t,0) = -e^{-t}\left[\int_0^t \left(\frac{l\partial i_0}{\partial t} + ri_0\right)dx\right]$$

$$= e^{-t}\int_{-t}^{0} [f_0(t+\lambda) + f_0'(t+\lambda)]I_0(\lambda)d\lambda. \quad (22)$$

If the expression (22) is inserted for $V_0(t, 0)$ in the equations derived from Kirchhoff's laws applied to the terminal network, an integro-differential equation for $f_0(t)$ will in general be obtained. In a sense, this equation corresponds to the integral equations which Carson[8] employs for the calculation of the current when certain indicial admittances are known.

Corresponding integral equations for the determination of the reflections at the terminals could also be formulated.

The integro-differential equation for $f_0(t)$ obtained by the use of expression (22) for $V_0(t, 0)$ is equivalent to Eq. (13) for $f_0(t)$. Since (14) is the solution of (13), it follows that (14) is also the solution of the corresponding integro-differential equation.[9] For values of t that are not too large, the solution (14) may be computed without great difficulty. For larger values of t, resort would have to be made to numerical integration as recommended by Carson.

[8] Reference 3, p. 142.

[9] This solution of the integro-differential equation appears to be equivalent to that given by H. T. Davis, "The Theory of the Volterra Integral Equation of the Second Kind," Indiana Univ. Studies (1930) and *The Theory of Linear Operators* (1936), p. 484.

The General Solution

The solutions discussed in the preceding are special solutions. Heaviside[10] in addition to the general solution for the unbounded line [see Eq. (3)] also obtained the general solution for the terminated line corresponding to $E(t)$ constant but with i and $\partial i/\partial t$ initially arbitrary functions of x. This Fourier series solution is of considerable complexity. Carson[11] has obtained a much simpler solution by an extension of the Heaviside calculus. It appears that the general solution may also be obtained in a tractable form by adding to the general solution of the unbounded line its reflections from the terminals. These latter can be computed by the methods given herein.

Acknowledgments

I am indebted to Mr. H. P. Hitchcock for evaluating the higher order terms in the series expansions, to Dr. L. S. Dederick, Mr. J. R. Lane and Mr. F. Reno for valuable suggestions and to Messrs. J. V. Holberton and J. H. Peck for the calculations on which Fig. 3 is based. I am grateful also to Dr. Carson of the Bell Telephone Laboratories for constructive criticisms.

[10] O. Heaviside, *Electrical Papers*, Vol. II. See also A. G. Webster, *Electricity and Magnetism*, p. 550.

[11] J. R. Carson, Bell Sys. Tech. J. 15, 340 (1936). See also B. van der Pol., Phil. Mag 7, 1153 (1929).

The Transition from Glow Discharge to Arc

HSU YUN FAN*
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received November 10, 1938)

INTRODUCTION

THE fundamental difference between glow and arc discharges lies in the mechanism of electron supply. The glow discharge requires a high cathode fall, because electrons must be liberated from the cathode by the comparatively inefficient agencies of ionic bombardment, photoelectric effect, or the impact of excited atoms generated in the gas. In the arc discharge, the cathode emits electrons by more efficient processes and the cathode fall is correspondingly lower.

If the cathode is made of a highly refractory material, the electron emission can be thermionic. Wehrli[1] has observed the gradual decline in the glow discharge cathode fall which accompanies the increase in temperature of cathodes of this kind. Arcs also occur, however, with cathodes of materials which, because of thermal disintegration, could never reach the high temperatures necessary for effective thermionic emission. Furthermore, glow discharges have been observed to change into arcs even when the cathodes were efficiently cooled, leading to the suggestion that the electrons are liberated from the cathode in this case by high electric fields. This hypothesis has, however, never been completely proved.

Transitions from glow to arc discharge can be produced either by an increase in current under constant gas pressure, or by an increase in pressure at constant current. The highest pressure under which glow discharges have been observed heretofore is of the order of one atmosphere.[2] This paper describes some further experiments on glow discharges in air, oxygen and nitrogen at one atmosphere pressure, and in hydrogen under pressures ranging from 1 to 13 atmospheres, using various combinations of electrode materials. These experiments have shown that glow discharges between copper electrodes may exist in hydrogen with currents as large as 14 amperes and, though not simultaneously, under pressures as high as 13 atmospheres; that with the other gases and electrode combinations tested the glow was less stable than in hydrogen; that continuous stability of the glow over long periods may not be expected even with small currents. Additional evidence was found for the hypothesis of electron emission in arcs under high surface field strengths, and characteristic changes of the electrodes in glow discharges have been observed due to sputtering and chemical reactions.

APPARATUS

The electrodes used consisted of disks approximately 12 mm in diameter and from 0.3 to 3.0 mm thick fixed to the ends of brass tubes. Very effective cooling was obtained by directing a jet of cold water against the center of the back of each electrode disk. The cooling was sufficiently effective that a tenfold increase over normal in the flow of water produced no change in the behavior of the discharge. The current for the discharge was supplied in most cases by two 2000-volt, one-ampere, d.c. generators which could be connected in series or parallel. For the largest currents one 550-volt and two 230-volt generators of large capacity were connected in series. In all cases the discharge was initiated by a high frequency, high voltage impulse, which was kept out of the generators by a suitable filter.

In the experiments at atmospheric pressure a chamber of about 3½ liters volume housed the electrodes. The gas was caused to stream through the chamber at a rate sufficient to carry off products of the discharge quickly. The hydrogen used at atmospheric pressure was prepared electrolytically, great care being taken to keep it

* The material included in this paper is taken from a thesis submitted to the Electrical Engineering Department of the Massachusetts Institute of Technology in May 1937 in partial fulfillment of the requirements for the degree of Doctor of Science. Publication has been delayed by the author's return to China, where he is now connected with Tsing Hua University.

[1] M. Wehrli, Helv. Phys. Acta **1**, 323 (1928).

[2] H. Thoma and L. Heer, Zeits. f. tech. Physik **13**, 464 (1932); **14**, 385 (1933); **15**, 186 (1934); R. Seeliger and K. Bock, Physik. Zeits. **34**, 767 (1933); O. Becken and R. Seeliger, Ann. d. Physik **24**, 609 (1935).

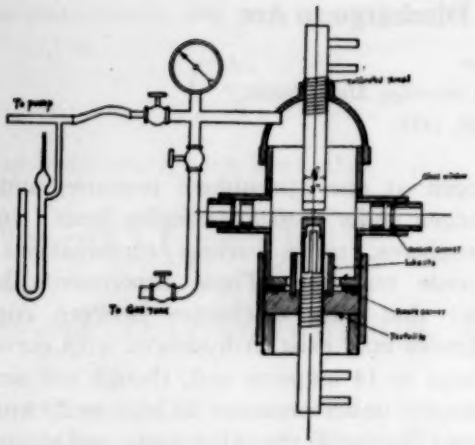
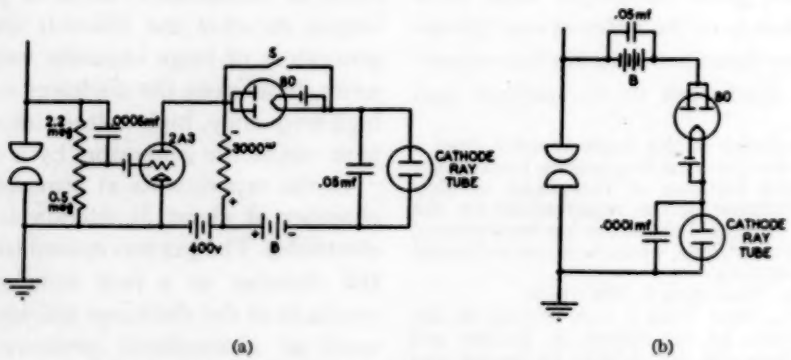


FIG. 1. Section of chamber used for studying discharges at pressures above atmospheric.

as pure as possible. The other gases were supplied from tanks, and the oxygen used was of ordinary commercial purity. The nitrogen and hydrogen were obtained especially, and had about one part in a thousand of impurity. The pressure chamber shown in Fig. 1 was used for the measurements with hydrogen above atmospheric pressure.

There are regions of current and pressure where the glow type of discharge is comparatively steady, and other different regions where the arc is steady. The experiments described in this paper pertain to the intermediate region where transitions from glow to arc take place. When the current is sufficiently large, temporary transitions from glow to arc can be detected. These produce a momentary increase in brightness of the discharge, particularly at the cathode, and are accompanied by a momentary drop in discharge voltage. When the discharge current is small the transition period is usually too brief to be observed by simple methods, and a fast detecting circuit must be used. The circuits of Fig. 2 were devised for this purpose, that of Fig. 2(a) having been developed first and serving for voltage drops of duration down to about 10^{-3} second, and that of Fig. 2(b), a modification of 2(a), serving for drops of duration as small as 10^{-6} second. Both circuits act on the principle that a condenser is charged when the voltage across the discharge falls, the charge being trapped by means of a thermionic rectifier when the voltage recovers. The presence and magnitude of the trapped charge is indicated by means of an electron oscillograph. In the circuit of Fig. 2(a) the deflection of the cathode-ray beam caused by a fall in discharge voltage was comparatively permanent, and it was necessary to return the beam to its original position by means of the switch *S*. In the circuit of Fig. 2(b) the internal leakage of the cathode-ray tube was sufficient, because of the small charge stored in the small capacitance across the tube, to return the beam in about one second to its equilibrium position. Observations with this circuit were facilitated by the simultaneous use of the horizontal sweep. The circuit of Fig. 2(b) gave a fairly accurate measure of the magnitude of the temporary change in voltage drop across the discharge, but the amplifier of Fig. 2(a) was not sufficiently carefully constructed to give more than qualitative information on the magnitude of the voltage changes. Decreases in glow discharge voltage drops were observed for all currents, and most of these decreases were in excess of 250 volts.



(a) (b)

FIG. 2. Circuits used in the study of voltage disturbances of short duration.

Procedure

The systems investigated are shown in Table I.

Air

In Fig. 3 the dotted curve shows the volt-ampere characteristic of the discharge in air at atmospheric pressure with copper electrodes. The glow becomes increasingly unstable as the current is increased, and the arc becomes increasingly unstable as the current is decreased. The voltage of the glow is constant over a wide range as is the rule for "normal" glow discharges, in which the surface of the cathode is not entirely covered by the negative glow. At very low currents the voltage increases sharply, but the curve could be followed over only a short range because of the limited voltage of the generators, and the need of using a large series stabilizing resistance.

In the lower ranges of current this discharge appears to the eye to be perfectly stable, a fact which has perhaps led some observers to postulate a lower limit of current below which transitions do not occur. It is true that the frequency of the transitions decreases with decreasing current, but with the fast detecting circuits used, transitions were observed down to the limit of current where further decrease became impractical because of the sharp increase in voltage (cf. Fig. 3). Substitution of platinum for copper in the electrodes did not change appreciably the frequency of the disturbances, which indicates

Table I.

Gas	Cathode	Anode	Pressure (atmos.)
Air	Cu	Cu, Pt	1
Air	Pt	Pt	1
O_2	Cu, Pt	Cu	1
N_2	Cu, Pt, Al	Cu	1
H_2	Cu, Al	Cu	1
H_2	Cu, Pt, Mo	Cu	1 to 13

that oxidation of the cathode is not the main cause of their occurrence.

In addition to the large voltage drops described above, small voltage drops (10–20 volts), following each other in rapid succession, were observed when current had been passed between the electrodes for a long period of time. At the higher currents these disturbances can be recognized immediately as being associated with the

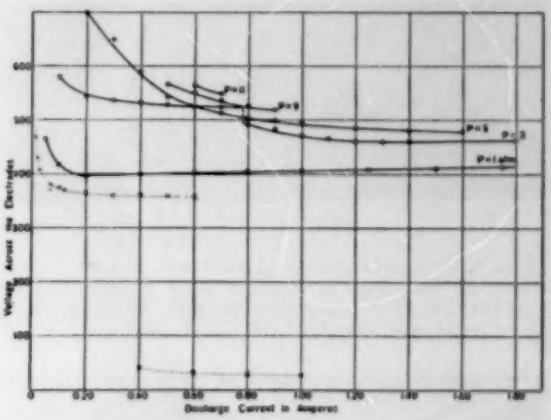


Fig. 3. Volt-ampere characteristics of a discharge in air at atmospheric pressure between water-cooled copper electrodes separated 1.0 mm (dotted curve), and of a glow discharge in hydrogen between water-cooled copper electrodes separated 0.9 mm at several pressures.

unsteady behavior of the anode spot and vaporization of the anode.

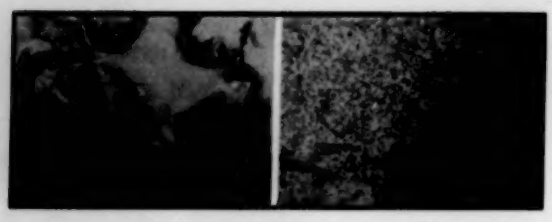
On the anode, the discharge leaves a trace consisting of a clean center surrounded by a black deposit. Once the anodic spot has moved out of the center it falls on the deposit, which begins to evaporate, thus causing further displacements of the spot and the voltage unsteadiness described above. The trace on the cathode shows a smooth spot corresponding in size to the negative glow. This spot, which is perfectly smooth even when enlarged 750 times, is surrounded by a deposit of fine black powder probably removed from the cathode by sputtering. Discharge traces on platinum electrodes do not differ essentially from those on copper electrodes.

Oxygen

Voltage disturbances are more frequent in oxygen than in air. The discharge traces on the electrodes have the same appearance, but the center of the cathode spot on platinum is rougher than in air.

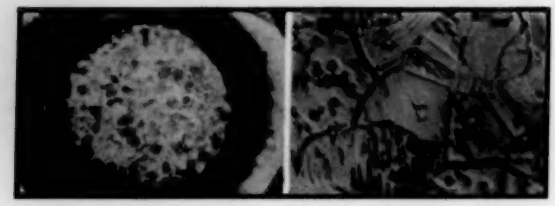
Nitrogen

Voltage disturbances are less frequent and less violent in nitrogen than in oxygen or air, and a comparatively stable glow can be maintained at much higher currents, sometimes up to two amperes. However, the discharge is restless, and to keep the anodic spot stationary the anode must be cone-shaped. Even when the anodic spot is



(a) (b)

FIG. 4. Photomicrographs of surface of copper cathode after glow discharge in nitrogen at 1 atmosphere. (a) magnification of 50 diameters showing irregular shape. (b) center of spot magnified 750 diameters.



(a) (b)

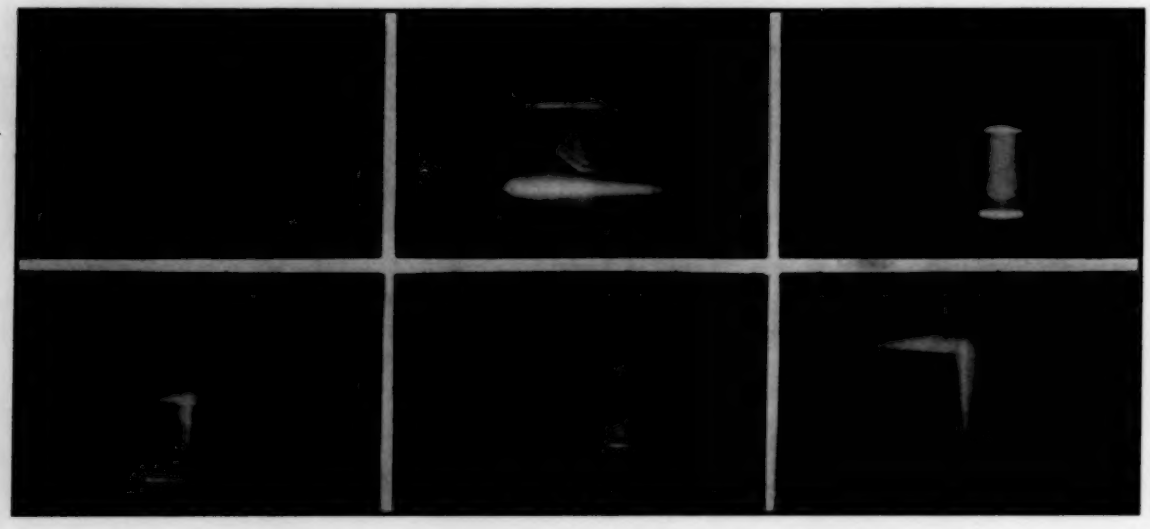
FIG. 5. Photomicrographs of surface of copper cathode after glow discharge in hydrogen at 5 atmospheres pressure. (a) magnification of 50 diameters. (b) center of spot magnified 750 diameters.

fixed on the top of the cone the irregularly shaped cathode glow refuses to stay still. The anode is not appreciably affected by the discharge, but Fig. 4(a) shows the irregular shape of the spot produced on the copper cathode and the destruction of the surface polish at its center. With aluminum cathodes (three mm thick) the discharge was much less stable than with copper cathodes. Disturbances were more frequent, and the discharge often broke down to a continuous arc running rapidly around the electrode. This arc quickly roughened and destroyed the polished cathode surface until it was no longer capable of supporting a glow discharge.

Hydrogen

The glow between copper electrodes is very stable in hydrogen, with no disturbances occurring at two amperes for several hours. In two tests at five amperes the glow broke down to a continuous arc after 40 and 25 minutes. In a test at ten amperes, the breakdown occurred after 15 minutes. To reach these large values without breakdown, the current must be increased slowly, and by this procedure the current was raised in one case to 14 amperes before the glow changed to an arc.

With cathodes of Pt (0.3 mm), Mo (0.375 mm) and Al (3 mm) disturbances were more frequent in hydrogen than with copper cathodes. The combination Pt-H_2 gave, however, less frequent disturbances than that Pt-N_2 combination. Re-



(a) (b) (c)

(d) (e) (f)

FIG. 6. Appearance of the discharge in hydrogen with copper electrodes separated 3 mm and at a variety of pressures and currents. (a) 1 atmosphere and 0.25 ampere, (b) 1 atmosphere and 2.0 amperes, (c) 2 atmospheres and 0.25 ampere, (d) 2 atmospheres and 1.0 ampere, (e) 4 atmospheres and 0.25 ampere, (f) 4 atmospheres and 1.0 ampere.

sults with molybdenum were similar to those with platinum; aluminum showed the same behavior in hydrogen as it did in nitrogen. The anodes were made of copper in all these experiments and they were not appreciably affected by the discharge.

The stability of the glow discharge in hydrogen with copper electrodes permitted increasing the pressure above one atmosphere, and the glow characteristics under different pressures are shown in Fig. 3. The region of comparatively

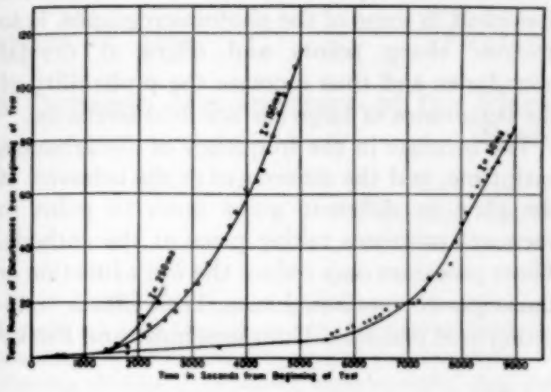


FIG. 7. Total number of disturbances from the beginning of a typical test (copper electrodes in air at atmospheric pressure) as a function of elapsed time from beginning of test.

stable glow becomes narrower with increasing pressure. The lower limit of this region, where the voltage begins to increase at low currents, shifts to higher currents, and at the same time the upper limit is shifted down by an increase in the frequency of disturbances. By using a very high voltage and a very large series resistance the glow can be stabilized with somewhat lower currents, but the limit is soon reached where it degenerates into a Townsend discharge.

Figure 5 shows the effects on the copper cathode of the discharge in hydrogen at a pressure of five atmospheres. The circular spot shown in the photograph of low magnification is surrounded by a circle of reddish deposit. The larger magnification shows how the crystal structure is exposed through sputtering.

Figure 6 shows the appearance of the glow in hydrogen at various pressures and currents. At the lower values of current and pressure the positive column shows well-defined striations. With increasing values the striations become diffused and eventually (Fig. 6(f)) the column concentrates abruptly into a narrow line, with a simultaneous drop in the voltage of the discharge of about 70 volts. It seems likely that this change results from temperature ionization in the column.

If a few percent of oxygen are added to hydrogen some additional striations are observed and the measuring circuit indicates frequent drops of voltage. The oxygen is soon cleaned up, as indicated by the disappearance of the additional striations, and the disturbances gradually disappear.

TIME EFFECTS

With any combination of gases and freshly polished electrodes the time interval before the first disturbance is usually longer than subsequent intervals between disturbances, and the frequency of disturbances gradually rises until it reaches a more or less constant value. The curves of Fig. 7 illustrate this behavior, though these curves must be considered as only qualitative, since the data on which they are based vary widely from trial to trial under apparently identical conditions. In these curves the total number of disturbances taking place from the beginning of the test up to any particular time is plotted as ordinate, with the time in question as abscissa. Consequently the slope of the curve in any region gives the frequency of occurrence of transitions. This is seen to reach a fairly constant value, with the time required to reach constancy less for the larger currents.

To determine whether the interval between the initiation of the discharge and the first transition was entirely random in length, a series of measurements of this interval was made, the electrode surfaces being carefully polished after each breakdown. The result of a typical series of one hundred tests, made in this case with copper electrodes in air at atmospheric pressure and a current of 0.31 ampere, is shown in Fig. 8. The block curve of this figure, showing the distribution in length of intervals between initiation of the discharge and the first disturbance, indicates that the length of this interval is not random, but has a most probable value, a sort of "mean formation time" during which the

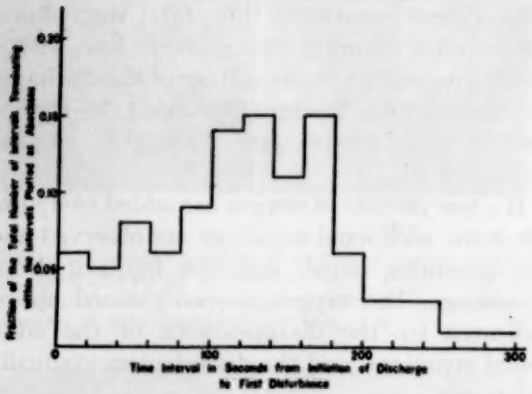


FIG. 8. Distribution in length of intervals between initiation of the discharge and first disturbance, using copper electrodes in air at atmospheric pressure, and a current of 0.31 ampere.

electrode surfaces undergo some kind of progressive change which finally disturbs the stability of the glow.

CURRENT DENSITY AT THE CATHODE

In all of the gases used except nitrogen the discharge formed a well-defined spot on the cathode. From the area of this spot and the total current, a cathode current density can be calculated, and the results of such calculations for discharges in air and in hydrogen are shown in Fig. 9. The current density is constant in air over the range studied, but in hydrogen it increases substantially in the region of small currents.

DISCUSSION

Calculation of the temperature distribution within the materials of the cathode, taking into account the effective cooling of the back surface, shows that the exposed surface could not reach the temperatures required to give appreciable thermionic emission. Since field emission is then the most likely source of cathode electrons during the intervals when the glow breaks down to an arc, the field strengths in the discharges studied were calculated in a manner similar to that used by Engel and Steenbeck.[3] The values found lie between 10^5 and 10^6 volts per cm indicating that field emission from small surface irregularities may be expected. Since the field strength at the cathode surface is determined by space charge, which is in turn dependent on current density, no lower limit of current for the occurrence of disturbances should exist. This is in accord with the experimental results. The increase in frequency of the disturbances with increasing current is to be expected, since the increased current spreads the cathode glow over a larger area, thus increasing the probability of the inclusion in this area of microscopic rough spots.

The effect of sputtering of the cathode, which is evident in some of the photomicrographs, is to uncover sharp points and edges at crystal boundaries and thus increase the probability of the occurrence of large surface field strengths.

The increase in the frequency of disturbances with time, and the differences in the behavior of the glow in different gases seem to point to chemical processes taking place at the cathode. These processes may reduce the work function or cause gas or vapor evolution, both effects facilitating field emission. Güntherschulze and Fricke[4]

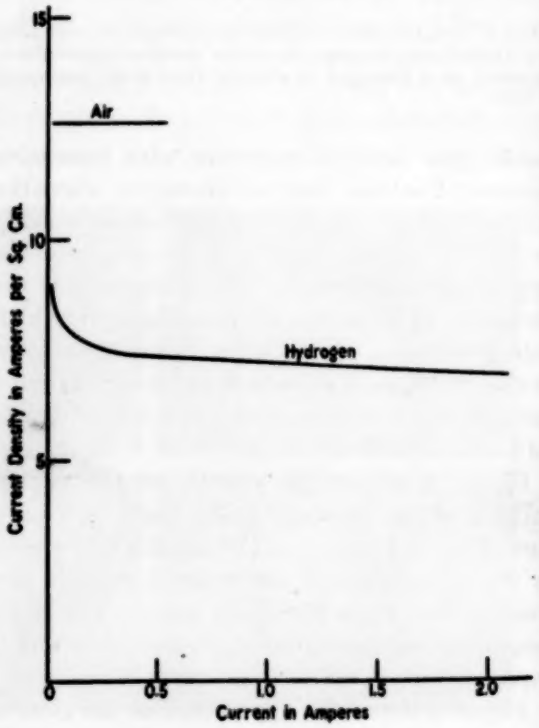


FIG. 9. Current density at the cathode surface as a function of total current in the discharge.

[3] A. von Engel and M. Steenbeck, *Elektrische Gasentladungen II.*

[4] A. Güntherschulze and H. Fricke, Zeits. f. Physik **86**, 451 (1933); H. Fricke, Zeits. f. Physik **92**, 728 (1934).

have observed that glow discharges without cathode-fall are obtained with cathodes made of semi-conductors covered with a thin layer of insulating material. Metal cathodes covered in the same way have a cathode fall but the discharge is accompanied by small sparks on the cathode surface. In a similar way insulating layers formed chemically on the cathode surfaces in the experiments described might give rise to the flashes observed at large currents.

The perfectly smooth surfaces of the cathode spots in air and oxygen and the restlessness of the negative glow in nitrogen are direct evidences of complicated phenomena on the cathode surface. More definite statements about the chemical effects cannot be made, however, until the disturbances have been studied under very pure conditions.

APRIL 15, 1939 PHYSICAL REVIEW VOLUME 55

A Dynamic Measurement of the Elastic, Electric and Piezoelectric Constants of Rochelle Salt

W. P. MASON
Bell Telephone Laboratories, New York, New York
(Received November 12, 1938)

The elastic, electric and piezoelectric constants of Rochelle salt have been measured at low field strengths by measuring the resonant frequencies and impedance of vibrating crystals. It is shown experimentally that the resonant and antiresonant frequencies of the crystal are both considerably below the natural mechanical resonant frequency of the crystal in disagreement with the usual derivation of the frequencies of a piezoelectric crystal. By assuming that the piezoelectric stress is proportional to the charge density on the electrodes rather than the potential gradient as usually assumed, theoretical frequencies are obtained which agree with those found experimentally. This theoretical derivation together with the measured frequencies supply values for the piezoelectric constants. The elastic constants measured dynamically show some differences from those measured statically. A large difference is found for the dynamically measured piezoelectric constants from those statically measured, which may be attributed to the finite relaxation time for the piezoelectric elements.

I. INTRODUCTION

THE static properties of Rochelle salt crystals have formed the subject for a number of measurements. W. Mandell[1] has measured the nine elastic constants of Rochelle salt by a static method, and a number of measurements have been made on the dielectric constants and the piezoelectric constants. However, very few measurements[2] have been made on the dynamic characteristics of Rochelle salt, which may differ considerably from the static characteristics due principally to the large relaxation time of the piezoelectric elements of the crystal. It is the purpose of this paper to present dynamic measurements on the elastic, dielectric and piezoelectric constants of Rochelle salt crystals. These indicate a considerable difference from those determined statically, particularly the piezoelectric and the dielectric constants. Because of the precision obtainable in frequency determination it is felt that the measurements made in this way should be more accurate than those obtained by other methods.

In measuring the elastic constants of a crystal, the frequency of resonance of some known mode of motion is measured electrically and the elastic constant is calculated from this measurement and the known density of the crystal. The usual method of deriving the equations of motion[3]

[1] "The Determination of the Elastic Modulii of the Piezo-Electric Crystal Rochelle Salt by a Statical Method," W. Mandell, Proc. Roy. Soc. London **116**, 623 (1927).

[2] One of the dynamic piezoelectric constants was recently measured by G. Mikhailov (Tech. Phys. U. S. S. R. **4**, 461 (1937)) using a different method. He does not obtain as high a piezoelectric constant as measured by this method.

[3] See for example *Quartz Resonators and Oscillators*, P. Vigoureux (H.M. Stationery Office, London), Chapter III, or *Piezo-electrizität Des Quarzes*, Adolf Schiebe (Theodor Steinkopff, 1938), p. 85.

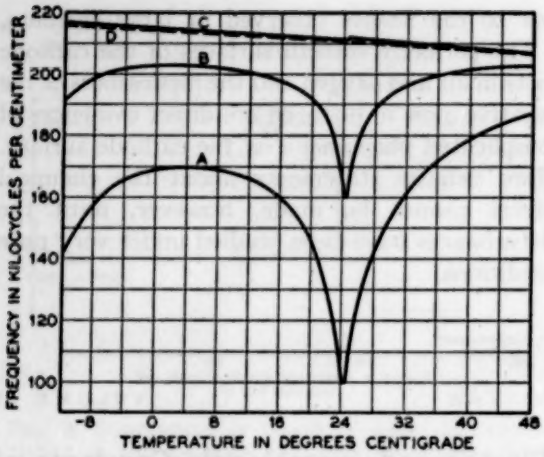


FIG. 1. Measured frequencies for a longitudinally vibrating Rochelle salt crystal measured as a function of the temperature. Curves *A* and *B* are, respectively, the measured resonance and antiresonance frequencies of a fully plated crystal. Curve *C* is the measured resonance (or antiresonance) frequency of a crystal lightly plated at the center or an unplated crystal in an air-gap holder. Curve *D* the calculated natural mechanical resonance of the fully plated crystal.

(which is based on the assumed proportionality between piezoelectric stress and applied potential gradient in the crystal) indicates that the resonant frequency of the crystal measured electrically coincides with the natural mechanical resonance of the crystal.[4] As is evident from a measurement of the resonant and antiresonant frequencies of a highly coupled crystal such as Rochelle salt, this deduction does not agree with experiment. The data on Fig. 1 show measured resonance curves for a longitudinally vibrating Rochelle salt crystal with its major surfaces perpendicular to the *a* or electric axis of the crystal, and the length of the crystal laying 45° from the *b* and *c* axes. The curves labeled *A* and *B* are, respectively, the electrical resonant frequencies (frequencies of smallest impedance) and the electrical antiresonant frequencies (frequencies of largest impedance) of a fully plated crystal plotted as a function of the temperature of the crystal. The dimensions of the crystal—which were chosen so that no secondary resonances were present near the main one—were length = 2.014 cm; width = 0.418 cm; thickness = 0.104 cm; and the measured frequencies were expressed in terms of a crystal one centimeter long. The curve labeled *C* is a measurement of the resonant frequency (or antiresonant frequency since the two are nearly identical) of the same crystal with the plating removed except for a small amount at the center. The curve *C* is also obtained for an unplated crystal in an air-gap holder with a large air gap. As is evident the curve *C* is considerably higher in frequency than either *A* or *B*.

If we assume that the piezoelectric stress is proportional to the applied field, it is evident that the curves *A* and *C* should be identical, for the piezoelectric driving force will be in phase with the applied potential irrespective of the amount of the crystal plated and hence in both cases the resonant frequency measured should agree with the natural mechanical resonant frequency of the crystal, a deduction that does not agree with experiment.

This question is considered in Appendix II and it is there shown that if the piezoelectric stress is taken proportional to the charge density on the electrodes, resonant and antiresonant frequencies are obtained which agree with the experimental results. From these theoretical results and the measured frequencies of Fig. 1, one of the piezoelectric constants of the crystal can be evaluated. It is shown also in Appendix II that if the driving plates and the crystal are separated by a large air gap, the electrical resonant and antiresonant frequencies coincide with the natural mechanical resonant frequency of the crystal and hence by measuring the electrical frequencies of the crystal in an air-gap holder, the true elastic constants of the crystal can be evaluated.

II. Measurements of the Nine Elastic Constants of Rochelle Salt

Rochelle salt belongs to the rhombic hemihedral class of crystals and possess three diagonal axes of symmetry which are identical with the three crystallographic axes. As a consequence of its class of symmetry it has nine elastic constants, and three piezoelectric constants all of which are shears for the three crystallographic axes.

If we associate the *a* axis with an *X* axis, the *b* axis with a *Y* axis, and the *c* with a *Z* axis, the

[4] For a plated crystal this deduction is obvious since for a plated crystal the applied potential is constant along the surface and hence the applied piezoelectric forces cancel out everywhere except at the two ends. The applied mechanical force on the ends of the crystal will be in phase with the applied potential and hence the electrical and mechanical resonances will occur at the same frequency.

elastic equations for a Rochelle salt crystal are

$$-x_x=s_{11}X_x+s_{12}Y_y+s_{13}Z_z; \quad -y_z=s_{44}Y_z,$$
$$-y_y=s_{12}X_x+s_{22}Y_y+s_{23}Z_z; \quad -z_x=s_{55}Z_x, \quad (1)$$
$$-z_z=s_{13}X_x+s_{23}Y_y+s_{33}Z_z; \quad -x_y=s_{66}X_y,$$

in which s_{11} etc. are the nine elastic compliances, x_x etc. the strains, and X_x etc. the stresses. We may also write this relationship in the form

$$-X_x=c_{11}x_x+c_{12}y_y+c_{13}z_z; \quad -Y_z=c_{44}y_z,$$
$$-Y_y=c_{12}x_x+c_{22}y_y+c_{23}z_z; \quad -Z_x=c_{55}z_x, \quad (1')$$
$$-Z_z=c_{13}x_x+c_{23}y_y+c_{33}z_z; \quad -X_y=c_{66}x_y,$$

where

$$\sigma c_{11}=\begin{vmatrix} s_{22}s_{23} \\ s_{23}s_{33} \end{vmatrix}; \quad \sigma c_{23}=\begin{vmatrix} s_{13}s_{23} \\ s_{11}s_{12} \end{vmatrix};$$

$$c_{44}=\frac{1}{s_{44}} \quad \text{where} \quad \sigma=\begin{vmatrix} s_{11}s_{12}s_{13} \\ s_{12}s_{22}s_{23} \\ s_{13}s_{23}s_{33} \end{vmatrix}.$$

The simplest way to measure the modulii of compliance s_{11}, s_{22}, s_{33} (which are the inverse of Young's modulii along the X, Y and Z axes) is to cut crystals having their lengths large compared to any other dimensions, and lying respectively along the X, Y and Z axes, then to measure the frequencies of resonance of such crystals. The value of the modulus of compliance is obtained from the formula

$$f=\frac{1}{2l}\left(\frac{1}{\rho s_{nn}}\right)^{\frac{1}{2}} \quad \text{or} \quad s_{nn}=\frac{1}{(2lf)^2\rho}, \quad (2)$$

where l is the length of the crystal, f is the frequency of mechanical resonance, and ρ the density of Rochelle salt which after careful measurement[5] was found to be

$$\rho=1.775. \quad (3)$$

Unfortunately, however, crystals cut along these three axes have no components of driving force to drive them in a longitudinal mode. However, if the crystal is cut at an angle to any of the three axes, the piezoelectric constant determining the longitudinal driving force is finite as shown in Appendix I, and a measurable resonance is obtained. Accordingly crystals were cut with their thin dimensions perpendicular to the X, Y and Z axes, respectively, with their long dimensions at various angles from the other crystallographic axes. These crystals and their frequencies measured in an air-gap holder are given on Table I. The frequencies were all measured at 30°C.

TABLE I.

(1) Direction of Applied Potential	(2) Angle From	(3) Dimensions in mm L	W	T	(4) Frequency of Resonance Cycles	(5) Frequency Constant kc per cm	(6) Value of s Constants
1 X	Y=22.5°	20.02	2.14	1.08	102,420	204.9	3.36×10⁻¹²
2 X	Y=45°	20.22	2.24	.98	104,110	210.5	3.16
3 X	Y=67.5°	19.72	2.26	1.06	105,670	208.4	3.23
4 Y	X=67.5°	19.86	2.16	1.06	76,340	151.6	6.14
5 Y	X=45°	20.06	2.08	1.02	61,176	122.7	9.31
6 Y	X=22.5°	19.94	2.14	.98	68,820	137.25	7.45
7 Z	X=67.5°	19.84	2.08	.96	100,810	200.0	3.525
8 Z	X=45°	19.65	2.12	.98	96,570	189.9	3.905
9 Z	X=22.5°	19.42	2.12	.98	89,320	173.2	4.69

In order to use these values it is necessary to know how Young's modulus varies for different angular cuts. The variation with orientation is worked out in Appendix I. For a crystal cut perpendicular to the a or X axis this becomes

$$s_{22}'=s_{22}\cos^4\theta+(2s_{23}+s_{44})\sin^2\theta\cos^2\theta+s_{33}\sin^4\theta, \quad (4)$$

where θ is the angle between the long dimension and the Y axis. The value of s_{22}' is given for three angles by the first three values of column 6, and hence the three unknown quantities s_{22}, $(2s_{23}+s_{44})$ and s_{33} can be solved for. We find

$$s_{22}=3.50\times10^{-12}; \quad s_{33}=3.32\times10^{-12}; \quad (2s_{23}+s_{44})=5.93\times10^{-12}. \quad (5)$$

For a rotation about the Y axis the expression for s_{11}' becomes

$$s_{11}'=s_{11}\cos^4\theta+(2s_{13}+s_{55})\sin^2\theta\cos^2\theta+s_{33}\sin^4\theta, \quad (6)$$

where θ is the angle measured from the X axis. The data for crystals 4, 5 and 6 are applicable to this case, and we find

$$s_{11}=5.23\times10^{-12}; \quad s_{33}=3.365\times10^{-12}; \quad (2s_{13}+s_{55})=28.6\times10^{-12}. \quad (7)$$

[5] This measurement was made by Mr. W. L. Bond.

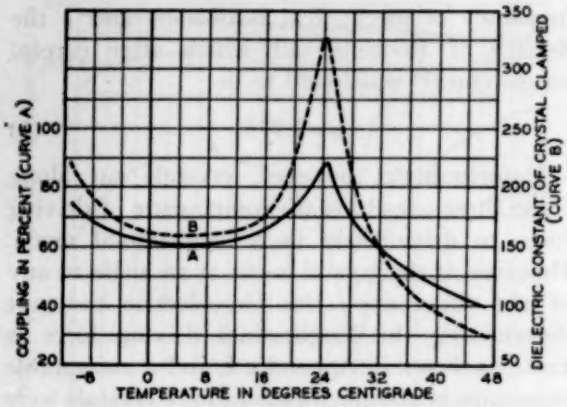


FIG. 2. Dynamic electric and coupling constants of an X cut Rochelle salt crystal. Curve A represents the electromechanical coupling of the crystal. Curve B represents the dielectric constant of the crystal clamped, plotted as a function of the temperature.

Similarly for a rotation about Z we find

$$s_{11}' = s_{11}\cos^4\theta + (2s_{12}+s_{66})\sin^2\theta\cos^2\theta + s_{22}\sin^4\theta \quad (8)$$

and from the data for crystals 7, 8 and 9 we find

$$s_{11}=5.14\times10^{-12};\quad s_{22}=3.495\times10^{-12};\quad (2s_{12}+s_{66})=7.02\times10^{-12}. \quad (9)$$

There are two independent values for s_{11}, s_{22} and s_{33} and a comparison shows that they agree within about one percent, which is the order of accuracy of the measurement. By these measurements we have evaluated three constants, and have three relations between the remaining six constants.

In order to evaluate the remaining constants it is necessary to measure the three shear coefficients s_{44}, s_{55}, s_{66}. It is shown in Appendix I, that if a crystal is cut with one edge of its major surface lying along one of the crystallographic axes and the normal to the major surface at an angle between the other two crystallographic axes, a piezoelectric constant results which sets up a high frequency shear wave in the crystal similar to the high frequency shear vibration excited in Y or AT cut quartz. The resonances of a large thin plate have been worked out by E. B. Christofel[6] who showed that there were three different values whose velocity of propagation could be obtained from the determinant

$$\begin{vmatrix} \lambda_{11}-\rho v^2, & \lambda_{12}, & \lambda_{13} \\ \lambda_{12}, & \lambda_{22}-\rho v^2, & \lambda_{23} \\ \lambda_{13}, & \lambda_{23}, & \lambda_{33}-\rho v^2 \end{vmatrix} = 0. \quad (10)$$

In this equation ρ is the density, v the velocity of propagation, and the λ's are related to the elastic constants of the crystal by the formulae

$$\begin{aligned}
\lambda_{11} &= c_{11}l^2+c_{66}m^2+c_{55}n^2+2c_{56}mn+2c_{15}nl+2c_{16}lm,\\
\lambda_{22} &= c_{66}l^2+c_{22}m^2+c_{44}n^2+2c_{23}mn+2c_{46}nl+2c_{26}lm,\\
\lambda_{33} &= c_{55}l^2+c_{44}m^2+c_{33}n^2+2c_{34}mn+2c_{35}nl+2c_{45}lm,\\
\lambda_{12} &= c_{16}l^2+c_{26}m^2+c_{45}n^2+(c_{46}+c_{25})mn+(c_{14}+c_{56})nl+(c_{12}+c_{66})lm,\\
\lambda_{13} &= c_{15}l^2+c_{46}m^2+c_{35}n^2+(c_{45}+c_{36})mn+(c_{13}+c_{55})nl+(c_{14}+c_{56})lm,\\
\lambda_{23} &= c_{56}l^2+c_{24}m^2+c_{34}n^2+(c_{44}+c_{25})mn+(c_{36}+c_{45})nl+(c_{25}+c_{46})lm,
\end{aligned} \quad (11)$$

where l, m and n are, respectively, the direction cosines between the normal to the surface which is the direction of propagation and the X, Y and Z axes. For Rochelle salt

$$c_{14}=c_{15}=c_{16}=c_{24}=c_{25}=c_{26}=c_{34}=c_{35}=c_{36}=c_{45}=c_{46}=c_{56}=0.$$

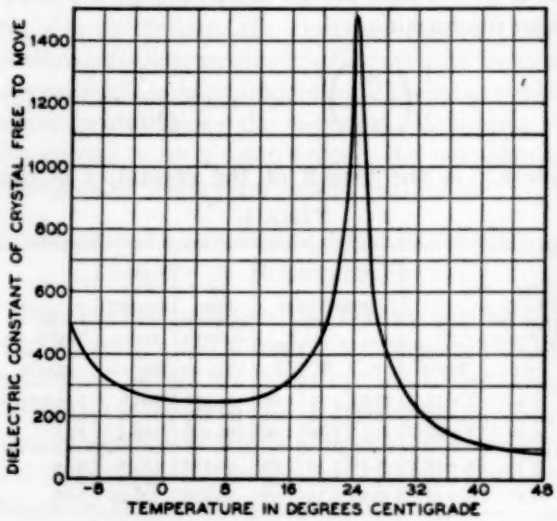


FIG. 3. Dielectric constant of an X cut Rochelle salt crystal free to move plotted as a function of the temperature.

[6] See A. E. H. Love's *Theory of Elasticity* (Cambridge University Press, 1934), fourth edition, p. 298.

For a crystal one edge of which lies along the X axis and the normal to the surface makes an angle θ with the Y axis $l=0$, $m=\cos\theta$, $n=\sin\theta$. Hence we find that

$$\lambda_{11}=c_{66}\cos^2\theta+c_{55}\sin^2\theta;\quad \lambda_{12}=\lambda_{13}=0;$$
$$\lambda_{23}=(c_{25}+c_{44})\sin\theta\cos\theta;$$
$$\lambda_{22}=c_{22}\cos^2\theta+c_{44}\sin^2\theta; \tag{12}$$
$$\lambda_{33}=c_{44}\cos^2\theta+c_{33}\sin^2\theta.$$

Since λ_{12} and λ_{13} are zero we have a shear mode of motion whose velocity is given by

$$v=(\lambda_{11}/\rho)^{\frac{1}{2}}$$ and whose frequency by

$$f=\frac{1}{2l_t}\left(\frac{c_{66}\cos^2\theta+c_{55}\sin^2\theta}{\rho}\right)^{\frac{1}{2}}. \tag{13}$$

Similarly for a crystal with one edge lying along the Y axis and the normal to the surface making an angle θ with the Z axis, we have a shear mode given by

$$f=\frac{1}{2l_t}\left(\frac{c_{44}\sin^2\theta+c_{66}\cos^2\theta}{\rho}\right)^{\frac{1}{2}}. \tag{14}$$

For a crystal lying along Z and whose normal makes an angle θ with the Y axis the frequency of resonance becomes

$$f=\frac{1}{2l_t}\left(\frac{c_{55}\cos^2\theta+c_{44}\sin^2\theta}{\rho}\right)^{\frac{1}{2}}. \tag{15}$$

These shear modes are the ones driven by the piezoelectric constants of the crystal.

In order to determine the constants c_{44}, c_{55} and c_{66} eight crystals were cut as shown on Table II and their resonant frequencies measured.

TABLE II.

Edge lies along	Angle of Normal from	Dimensions in mm			Frequency Constant kc per mm	Value of c constant	No. of Crystal
		L	W	T			
X	22.5° from Y	19.98	19.98	2.05	1129	7.51×10^{10}	1
X	45° from Y	29.24	29.24	1.92	945.2	6.34×10^{10}	2
X	67.5° from Y	24.00	23.98	2.36	755	4.06×10^{10}	3
Y	22.5° from X	25.44	25.44	2.28	1197	10.18×10^{10}	4
Y	45° from X	26.58	26.59	1.37	1265	11.39×10^{10}	5
Y	67.5° from X	20.66	20.64	1.72	1299	11.99×10^{10}	6
Z	22.5° from X	25.44	25.46	2.06	799.2	4.54×10^{10}	7
Z	45° from X	27.56	27.54	2.26	1052.0	7.87×10^{10}	8

Solving for c_{55} and c_{66} using the first three crystals of this table and Eq. (13) we have

$$c_{55}=2.96\times10^{10};\quad c_{66}=10.02\times10^{10}.$$

Using Eq. (14) and crystals 4, 5 and 6, we find

$$c_{44}=12.43\times10^{10};\quad c_{66}=9.91\times10^{10}.$$

Using Eq. (15) and crystals 7 and 8, we find

$$c_{44}=12.61\times10^{10};\quad c_{55}=3.13\times10^{10}.$$

Averaging these results we have for the three shear coefficients

$$c_{44}=12.52\times10^{10};\quad s_{44}=1/c_{44}=7.98\times10^{-12};$$
$$c_{55}=3.04\times10^{10};\quad s_{55}=1/c_{55}=32.8\times10^{-12};$$
$$c_{66}=9.96\times10^{10};\quad s_{66}=1/c_{66}=10.08\times10^{-12}.$$

Inserting these values in the relations given previously we find for the nine elastic constants the results shown on Table III. These are compared with the static results obtained by Mandell.

III. Measurements of Dielectric, Coupling and Piezoelectric Constants of Rochelle Salt

The three piezoelectric constants and the three dielectric constants along the three crystallographic axes have been measured by measuring the resonances and the capacitances of crystals cut with their major surfaces perpendicular to these three axes. This work was done at field strengths below 40 volts per centimeter, where

TABLE III.

	s_{11}	s_{22}	s_{33}	s_{44}	s_{55}	s_{66}	s_{12}	s_{13}	s_{23}
Coefficient by Dynamic Method	5.18×10^{-12}	3.495	3.34	7.98	32.8	10.08	−1.53	−2.11	−1.03
Coefficient Measured by Mandell	4.59×10^{-12}	3.205	2.815	6.06	30.6	8.02	− .795	−2.18	+1.69

	c_{11}	c_{22}	c_{33}	c_{44}	c_{55}	c_{66}	c_{12}	c_{13}	c_{23}
Coefficient by Dynamic Method	42.5×10^{10}	51.5	62.9	12.52	3.04	9.96	+29.6	+35.7	+34.25
Coefficient Measured by Mandell	34.7×10^{10}	47.3	80.6	16.35	3.24	12.42	− 6.05	+31.6	−34.4

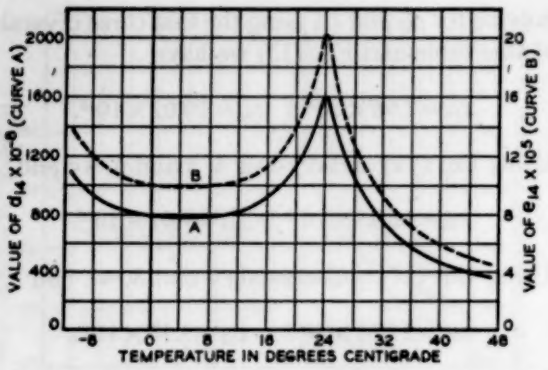


FIG. 4. Dynamic piezoelectric modulii of X cut Rochelle salt. Curve A is the piezoelectric constant d_{14} while curve B is the piezoelectric constant e_{14} plotted as a function of the temperature.

the expansion is a linear function of the applied field. The frequencies used were in the order of 80 kilocycles. Since it has been shown previously[7] that the constants are nearly independent of the frequency, above 100 cycles, the results can be used at other frequencies.

Figure 1 shows the measured resonance and antiresonance frequency of an X cut crystal with its length lying 45° from the Y and Z axes. With the aid of Fig. 7 of Appendix II, the electromechanical coupling defined by the Eq. (54) of Appendix II

$$k_x=\frac{d_{14}}{2}\left(\frac{4\pi}{Ks_{22}'}\right)^{\frac{1}{2}} \tag{17}$$

can be evaluated as a function of temperature and is plotted on Fig. 2, curve A. In Eq. (17) d_{14} is the piezoelectric constant, s_{22}' the inverse of Young's modulus along the length of the crystal and K the dielectric constant of the crystal when it is clamped or prevented from moving. In order to evaluate the piezoelectric constant d_{14} it is necessary to know the dielectric constant K. In order to determine this constant crystals were cut small enough to have no resonance near the frequency region of the resonances of the crystal of Fig. 1. The capacitance of the crystal free to vibrate was then measured over a temperature range and the dielectric constant resulting is shown on Fig. 3. It is shown in Appendix II, that the clamped dielectric constant is equal to the free dielectric constant divided by $(1-k^2)$. Hence from the data on Figs. 2 and 3 the dielectric constant K can be evaluated and it is plotted as curve B of Fig. 2. The value of s_{22}' for this crystal over a temperature range can be calculated from curve C of Fig. 1, hence from Eq. (17) the value of d_{14} can be evaluated and is shown as curve A of Fig. 4. Another piezoelectric constant e_{14} is sometimes used. This is related to d_{14} by

$$d_{14}/s_{44}=e_{14}. \tag{18}$$

The value of e_{14} is plotted as curve B of Fig. 4. The value of the piezoelectric constant obtained this way is smaller than that obtained statically, and moreover it shows a maximum at the Curie transformation point. Apparently the difference is due to the finite time of relaxation of the piezoelectric elements which do not allow them to attain their static value for an applied alternating voltage.

The values of the piezoelectric and dielectric constants along the other two axes were also measured and were found not to vary much with temperature. At 30°C, the Y constants were found to be

$$k_y=0.279;\quad K_y=12.5;\quad d_{25}=169\times10^{-8};\quad e_{25}=d_{25}/s_{55}=5.15\times10^{4}. \tag{19}$$

For the Z axis the results were

$$k_z=0.111;\quad K_z=10.2;\quad d_{36}=39.4\times10^{-8};\quad e_{36}=d_{36}/s_{66}=3.91\times10^{4}. \tag{20}$$

It will be noted that a dynamic method does not determine the sign of the piezoelectric constant but only its magnitude. From static measurements d_{25} is of opposite sign to d_{14} and d_{36}.

APPENDIX I

Constants of rotated rochelle salt crystals

The elastic equations of a Rochelle salt crystal are given by Eqs. (1) and (1′) when the a axis of the crystal is designated the X axis, the b the Y axis and the c the Z axis. The direct piezoelectric equations are

$$\begin{aligned} P_x&=d_{14}Y_z; \quad & P_y&=d_{25}Z_x; \quad & P_z&=d_{36}X_y, \\ -P_x&=e_{14}y_z; \quad & -P_y&=e_{25}z_x; \quad & -P_z&=e_{36}x_y, \end{aligned} \tag{21}$$

where P_x, P_y, P_z are the polarizations (charge per

[7] O. Norgorden, Phys. Rev. 49, 820 (1936).

unit area) along the X, Y and Z axes, respectively. The usual method of writing the converse piezoelectric equations is

$$\begin{aligned} -Y_x &= e_{14}E_x; \quad -Z_x = e_{25}E_y; \quad -X_y = e_{36}E_z, \\ y_z &= d_{14}E_x; \quad z_x = d_{25}E_y; \quad x_y = d_{36}E_z, \end{aligned} \tag{22}$$

where E_x, E_y, E_z are the potential gradients along the X, Y and Z axes.

In determining the constants of the crystal it is desirable to obtain cuts along other directions than the three crystallographic axes and it is desirable to obtain the elastic and piezoelectric constants for these other cuts. These cuts are determined by a new set of axes X', Y', Z' related to the old crystallographic axes by the direction cosines l_1 to n_3 expressed by the equation

$$\begin{array}{cccc} & X & Y & Z \\ X' & l_1 & m_1 & n_1 \\ Y' & l_2 & m_2 & n_2 \\ Z' & l_3 & m_3 & n_3. \end{array} \tag{23}$$

The transformations of stress as given by Love[8] are

$$\begin{aligned} X_x' &= l_1^2X_x + m_1^2Y_y + n_1^2Z_z + 2m_1n_1Y_z \\ &\quad + 2n_1l_1Z_x + 2l_1m_1X_y \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ X_y' &= l_1l_2X_x + m_1m_2Y_y + n_1n_2Z_z \\ &\quad + (m_1n_2 + n_1m_2)Y_z \\ &\quad + (n_1l_2 + n_2l_1)Z_x + (l_1m_2 + l_2m_1)X_y, \end{aligned} \tag{24}$$

where the primes denote the stresses referred to the new axes. Conversely we have

$$\begin{aligned} X_x &= l_1^2X_x' + l_2^2Y_y' + l_3^2Z_z' + 2l_2l_3Y_z' \\ &\quad + 2l_3l_1Z_x' + 2l_1l_2X_y' \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ X_y &= l_1m_1X_x' + l_2m_2Y_y' + l_3m_3Z_z' \\ &\quad + (l_2m_3 + l_3m_2)Y_z' + (l_3m_1 + l_1m_3)Z_x' \\ &\quad + (l_1m_2 + l_2m_1)X_y'. \end{aligned} \tag{25}$$

Similarly the transformation for the strain becomes[9]

$$\begin{aligned} x_x' &= l_1^2x_x + m_1^2y_y + n_1^2z_z + m_1n_1y_z \\ &\quad + n_1l_1z_x + l_1n_1x_y \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_y' &= 2l_1l_2x_x + 2m_1m_2y_y + 2n_1n_2z_z \\ &\quad + (m_1n_2 + m_2n_1)y_z + (n_1l_2 + n_2l_1)z_x \\ &\quad + (l_1m_2 + l_2m_1)x_y \end{aligned} \tag{26}$$

and conversely

$$\begin{aligned} x_x &= l_1^2x_x' + l_2^2y_y' + l_3^2z_z' + l_2l_3y_z' \\ &\quad + l_3l_1z_x' + l_1l_2x_y' \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_y &= 2l_1m_1x_x' + 2l_2m_2y_y' + 2l_3m_3z_z' \\ &\quad + (l_2m_3 + l_3m_2)y_z' + (l_3m_1 + l_1m_3)z_x' \\ &\quad + (l_1m_2 + l_2m_1)x_y'. \end{aligned} \tag{27}$$

Polarization, being a vector, transforms according to the equations

$$\left.\begin{aligned} P_x &= l_1P_x' + l_2P_y' + l_3P_z', \\ P_y &= m_1P_x' + m_2P_y' + m_3P_z', \\ P_z &= n_1P_x' + n_2P_y' + n_3P_z', \end{aligned}\right\} \tag{28}$$

and conversely

$$\left.\begin{aligned} P_x' &= l_1P_x + m_1P_y + n_1P_z, \\ P_y' &= l_2P_x + m_2P_y + n_2P_z, \\ P_z' &= l_3P_x + m_3P_y + n_3P_z. \end{aligned}\right\}$$

We now consider the simple case of a section whose X axis remains unchanged but whose other axes are rotated by an angle θ with respect to the crystallographic axes. For this case

$$\begin{aligned} l_1 = 1; \quad m_2 = n_3 = \cos\theta; \quad n_2 = -m_3 = \sin\theta; \\ l_2 = l_3 = m_1 = n_1 = 0. \end{aligned} \tag{29}$$

Then to determine the piezoelectric constants for a section of this kind we have

$$\begin{aligned} P_z' &= P_z = e_{14}y_z \\ &= e_{14}(y_y' \sin 2\theta - z_x' \sin 2\theta + y_z' \cos 2\theta). \end{aligned} \tag{30}$$

In the rotated crystal this would give a constant e_{12}' relating an expansion along the Y' axis with the X' polarization equal to $e_{14} \sin 2\theta$. Working out the polarization along the other two axes we find the constants

$$\begin{aligned} e_{12}' &= e_{14} \sin 2\theta; \quad e_{13}' = -e_{14} \sin 2\theta; \\ e_{14}' &= e_{14} \cos 2\theta; \\ e_{25}' &= [e_{25} \cos^2\theta - e_{36} \sin^2\theta]; \\ e_{26}' &= (e_{25} + e_{36}) \sin\theta \cos\theta; \\ e_{35}' &= -(e_{25} + e_{36}) \sin\theta \cos\theta; \\ e_{36}' &= e_{36} \cos^2\theta - e_{25} \sin^2\theta. \end{aligned} \tag{31}$$

[8] Reference 6, p. 80.
[9] Reference 6, p. 42.

These equations show that for a crystal cut perpendicular to the X axis and with the length at an angle between the Y and Z axes a longitudinal vibration can be set up by the constant e_{12}' or e_{13}'. This vibration will have its maximum driving force when $\theta=45°$. These equations show also that if one axis of the principal plane lies along X and the normal to the principal plane makes an angle θ with the Y crystallographic axis, a high frequency shearing mode is set up by the constant e_{26}' and this reaches its maximum when $\theta=45°$.

The d constants transform similarly to the e constants, except for d_{12}' and d_{13}' which are equal to

$$d_{12}'=\tfrac{1}{2}d_{14}\sin 2\theta;\quad d_{13}'=-\tfrac{1}{2}d_{14}\sin 2\theta. \quad (32)$$

If we express the elastic equations for the rotated axes it is readily shown that

$$\begin{aligned}
s_{11}'&=s_{11},\\
s_{22}'&=s_{22}\cos^4\theta+(2s_{23}+s_{44})\sin^2\theta\cos^2\theta\\
&\qquad+s_{33}\sin^4\theta,\\
s_{33}'&=s_{33}\cos^4\theta+(2s_{23}+s_{44})\sin^2\theta\cos^2\theta\\
&\qquad+s_{22}\sin^4\theta,\\
s_{44}'&=s_{44}\cos^2 2\theta+(s_{22}+s_{33}-2s_{23})\sin^2 2\theta,\\
s_{55}'&=s_{55}\cos^2\theta+s_{66}\sin^2\theta,\\
s_{66}'&=s_{66}\cos^2\theta+s_{55}\sin^2\theta,\\
s_{12}'&=s_{12}\cos^2\theta+s_{13}\sin^2\theta,\\
s_{13}'&=s_{13}\cos^2\theta+s_{12}\sin^2\theta,\\
s_{14}'&=(s_{13}-s_{12})\sin 2\theta,\\
s_{23}'&=s_{23}(\cos^4\theta+\sin^4\theta)\\
&\qquad+(s_{22}+s_{33}-s_{44})\sin^2\theta\cos^2\theta,\\
s_{24}'&=\sin^2 2\theta[(s_{23}+s_{44}/2-s_{22})\cos^2\theta\\
&\qquad+(s_{33}-s_{23}-s_{44}/2)\sin^2\theta],\\
s_{34}'&=\sin^2 2\theta[(s_{33}-s_{23}-s_{44}/2)\cos^2\theta\\
&\qquad+(s_{23}-s_{22}+s_{44}/2)\sin^2\theta],\\
s_{56}'&=(s_{55}-s_{66})\sin\theta\cos\theta.
\end{aligned} \quad (33)$$

For crystals rotated about the Y or Z axis, the constants can be obtained from these results by simply permuting the subscripts. Thus for any

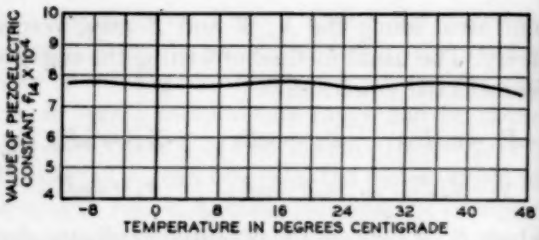


FIG. 5. A plot of the piezoelectric constant f_{14} as a function of the temperature.

subscript 1, 2 or 3, 1 is added for a rotation about Y and 2 for a rotation about Z with the understanding that $3+1=1$, i.e., the subscripts rotate among themselves. The same applies to the subscripts 4, 5 and 6. For example for a rotation about Y, we have

$$\begin{aligned}
s_{22}'&=s_{22},\\
s_{33}'&=s_{33}\cos^4\theta+(2s_{13}+s_{55})\sin^2\theta\cos^2\theta\\
&\qquad+s_{11}\sin^4\theta, \quad (34)\\
s_{11}'&=s_{11}\cos^4\theta+(2s_{13}+s_{55})\sin^2\theta\cos^2\theta\\
&\qquad+s_{33}\sin^4\theta, \text{ etc.}
\end{aligned}$$

For a rotation about Z

$$s_{11}'=s_{11}\cos^4\theta+(2s_{12}+s_{66})\sin^2\theta\cos^2\theta+s_{22}\sin^4\theta, \text{ etc.} \quad (35)$$

We require in the text also a relation between the two piezoelectric constants e_{14} and d_{14}. From Eq. (22) we have

$$Y_x/y_x=-e_{14}/d_{14}.$$

From Eq. (1)

$$Y_x/y_x=-1/s_{44}.$$

Hence $$d_{14}=e_{14}s_{44}. \quad (36)$$

Similarly $$d_{25}=e_{25}s_{55};\quad d_{36}=e_{36}s_{66}. \quad (37)$$

It is shown in Appendix II that in order to agree with experiment it is necessary to assume that the piezoelectric force is proportional to the charge density on the electrodes. This necessitates writing the converse piezoelectric equation in the form

$$-Y_x=f_{14}D_x;\quad -Z_x=f_{25}D_y;\quad -X_y=f_{36}D_z, \quad (38)$$

where $f_{14}=4\pi e_{14}/K_1$; $f_{25}=4\pi e_{25}/K_2$;

$$f_{36}=4\pi e_{36}/K_3$$

and D_x, D_y and D_z are the charge densities on the electrodes perpendicular to the three axes. In these equations K_1, K_2 and K_3 are the clamped dielectric constants along the X, Y and Z axes, respectively. For the most general piezoelectric crystal we can write the inverse piezoelectric equations in the form

$$\begin{aligned}
-X_x &= f_{11}D_x + f_{21}D_y + f_{31}D_z;\\
-Y_y &= f_{12}D_x + f_{22}D_y + f_{32}D_z;\\
-Z_z &= f_{13}D_x + f_{23}D_y + f_{33}D_z;\\
-Y_z &= f_{14}D_x + f_{24}D_y + f_{34}D_z;\\
-Z_x &= f_{15}D_x + f_{25}D_y + f_{35}D_z;\\
-X_y &= f_{16}D_x + f_{26}D_y + f_{36}D_z,
\end{aligned} \tag{39}$$

where $f_{ij} = 4\pi e_{ij}/K_i$.

The f constants for a rotated crystal can be obtained in the same manner as the d or e constants. For example for a rotation about the X axis, the f constants transform similarly to the e constants of Eq. (31). For Rochelle salt the f_{14} constant as a function of temperature is shown plotted on Fig. 5. The f_{25} and f_{36} constants have the values

$$f_{25} = 5.17\times10^4; \quad f_{36} = 4.81\times10^4,$$

which are nearly independent of temperature. The f constants are much more independent of temperature and orientation than the other piezoelectric constants.

It is readily shown by reference to equations in Appendix II that for static conditions the direct and converse piezoelectric equations take the form

$$\begin{aligned}
-P_x &= e_{14}y_x; \quad -P_y = e_{25}z_x; \quad -P_z = e_{36}x_y,\\
P_x &= d_{14}Y_x; \quad P_y = d_{25}Z_x; \quad P_z = d_{36}X_y
\end{aligned} \tag{21}$$

and

$$-Y_x = f_{14}D_x; \quad -Z_x = f_{25}D_y; \quad -X_y = f_{36}D_z,$$

$$y_x = \frac{4\pi}{K_1}d_{14}D_x = \frac{d_{14}E_x}{1-k_x^2}; \quad z_x = \frac{d_{25}}{1-k_y^2}E_y; \tag{22}'$$

$$x_y = \frac{d_{36}}{1-k_z^2}E_z.$$

We note that if the usual piezoelectric Eqs. (21) and (22) are assumed valid, a larger constant d_{14} will be measured by using the converse effect than by using the direct effect if the present theory is correct. This discrepancy has been noticed experimentally,[10] but was ascribed to a lack of reversibility in Rochelle salt. On the other hand if we express the expansion and charge on the plates in terms of the applied potential and force, it is readily shown that the method of expressing the piezoelectric equations used here satisfies the reversibility law, and hence the experimental discrepancy noted receives an explanation in terms of this theory.

For dynamic conditions only the first equations of (21) and (22)′ are valid.

Appendix II

Derivation of resonant frequencies of a piezoelectric crystal

As pointed out in the introduction, the measured resonant frequencies of a longitudinally vibrating Rochelle salt crystal do not agree with those given by the usual theoretical derivation since the measured resonance and antiresonance frequencies of the plated crystal are both considerably lower than the natural resonant frequency of the crystal measured by mechanical means. It is the purpose of this appendix to give a derivation of the equation of motion of a piezoelectric crystal, whose solution agrees with the experimental results.

Let us consider a small section of a plated piezoelectric crystal cut perpendicular to the X axis and with its length 45° from the Y and Z

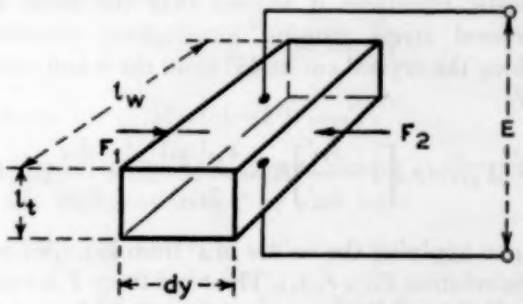


FIG. 6. A small section of a longitudinally vibrating Rochelle salt crystal.

[10] See "A Survey of Piezoelectricity," W. G. Cady, Am. Phys. Teacher 6, 237 (1938).

axes, as shown on Fig. 6. The width of the crystal is designated by l_w, the thickness by l_t and the elementary length along the Y' axis, the direction of vibration, by dy. If we denote the displacement of the section from its equilibrium position by ξ, then by Newton's law of motion

$$\rho l_w l_t dy(\partial^2\xi/\partial t^2) = +F_1 - F_2 = -(\partial F/\partial y)dy. \quad (40)$$

Here F_1 is the compressional force exerted normal to the left surface and F_2 that normal to the right-hand surface in the direction of increasing Y'. The difference between F_1 and F_2 for the small distance dy is the derivative of F by y as indicated on the right-hand side of (40). In this equation ρ is the density of the crystal and t the time.

We next consider the force strain equation. The usual method of deriving it is as follows: If we apply a potential gradient E_x along the thickness, the piezoelectric effect will apply a force per unit area equal to

$$-Y_y' = eE_x. \quad (41)$$

This constant e is not the constant $e_{12}' = e_{14}$ of Eq. (31) of Appendix I, but the constant $d_{14}/2s_{22}'$ where s_{22}' is the inverse of Young's modulus along the length of the crystal and d_{14} the usual shear piezoelectric modulus. This fact was first pointed out by Cady[11] for quartz, and for this cut of Rochelle salt can be shown as follows. From Eq. (31) of Appendix I, we see that a field applied along the X axis excites a Y_y' and a Z_z' stress and the last through the elastic cross constants (Poisson's ratio) makes contributions to Y_y'. From the fundamental piezoelectric and elastic equations it follows that the total impressed stress causing longitudinal vibration along the crystal cut at 45° from the b and c axes is

$$-(Y_y)_1 = e_{14}\left[1 - \frac{s_{23}'}{s_{22}'}\right]E_x = \frac{e_{14}(s_{44})}{2s_{22}'}E_x = \frac{d_{14}}{2s_{22}'}E_x \quad (42)$$

upon applying the values of s' from Eq. (33) and the relation $d_{14} = e_{14}s_{44}$. The total force F is equal to $Y_y(l_w l_t)$ while the potential gradient E_x is equal to the potential E applied divided by l_t. Hence

$$-\frac{F}{l_w l_t} = \frac{d_{14}}{2s_{22}'}\frac{E}{l_t} \quad \text{or} \quad -F = \frac{d_{14}}{2s_{22}'}l_w E. \quad (43)$$

This is the force applied by the piezoelectric effect if no motion takes place. If motion takes place an additional mechanical force is applied to the section equal to

$$-\frac{l_t l_w}{s_{22}' dy}[\xi_2 - \xi_1] = -\frac{l_t l_w}{s_{22}'}\frac{\partial \xi}{\partial y}. \quad (44)$$

Hence the total force will be the sum of the piezoelectric force and the elastic force or

$$-F = \frac{l_t l_w}{s_{22}'}\frac{\partial \xi}{\partial y} + \frac{d_{14}}{2s_{22}'}El_w. \quad (45)$$

Since E the applied potential is constant for a plated crystal and is not a function of y, Eqs. (40) and (45) could be combined into

$$\frac{1}{\rho s_{22}'}\frac{\partial^2\xi}{\partial y^2} = \frac{\partial^2\xi}{dt^2} \quad (46)$$

and the electrical resonant frequency of the crystal would coincide with the natural mechanical resonant frequency of the crystal, which does not agree with the experimental results of Fig. 1.

To get an equation which agrees with experiment it is necessary to assume that the piezoelectric force of Eq. (42) is proportional to the charge density on the electrodes rather than the potential gradient. For static conditions it is obvious that Eq. (43) can be expressed in the form

$$-F = d_{14}l_w dQ/2s_{22}'C \text{ since } dQ = CE, \quad (47)$$

where dQ is the charge on the electrode elementary area $l_w dy$, and C is the capacitance of this area which is equal in c.g.s. units to

$$C = l_w dy K/4\pi l_t. \quad (48)$$

In this formula K the dielectric constant has to be the dielectric constant with the crystal clamped and prevented from moving since the force Eq. (43) is valid only for this condition.

[11] W. G. Cady, "The Piezoelectric Resonator and the Effect of Electrode Spacing Upon Frequency," Physics 7, 245 (1936).

Then Eq. (45) becomes

$$-F=\frac{l_t l_w}{s_{22}'}\frac{\partial \xi}{\partial y}+\frac{d_{14}}{2s_{22}'}\frac{4\pi l_t}{K}\frac{dQ}{dy}, \tag{49}$$

where dQ now denotes the total charge on the surface. For dynamic conditions the total derivative will be replaced by the partial so that the equation can be written

$$-F=\frac{l_t l_w}{s_{22}'}\left[\frac{\partial \xi}{\partial y}+\frac{2\pi d_{14}}{l_w K}\frac{\partial Q}{\partial y}\right]. \tag{50}$$

One farther equation connects the applied voltage to the charge generated on the surface. The total charge on the surface consists of two parts, the charge placed on the surface by the applied potential, and the charge placed on by the piezoelectric effect when the crystal expands. This total charge will then be

$$dQ=\frac{El_w dyK}{4\pi l_t}-l_w dy e_{14}(y_y'-z_z') \tag{51}$$

by virtue of the piezoelectric relation (30) of Appendix I. But for a longitudinal vibration along Y', $z_z'=(s_{23}'/s_{22}')y_y'$ since s_{23}'/s_{22}' is Poisson's ratio by virtue of Eq. (1). Eq. (51) then becomes

$$dQ=\frac{l_w dyK}{4\pi}\left[\frac{E}{l_t}-\frac{4\pi d_{14}}{2s_{22}'K}\frac{\partial \xi}{\partial y}\right], \tag{52}$$

taking account of the relations given in Eq. (43), and the fact that $y_y'=\partial\xi/\partial y$.

For an unconstrained crystal F of Eq. (49) is zero so that

$$\frac{\partial \xi}{\partial y}=-\frac{4\pi d_{14}}{2l_w K}\frac{dQ}{dy}.$$

Substituting in (52) we find

$$C_F=\frac{dQ}{E}=\frac{Kl_w dy}{4\pi l_t(1-k^2)}, \tag{53}$$

where the coupling k is defined by the equation

$$k=\frac{d_{14}}{2}\left(\frac{4\pi}{Ks_{22}'}\right)^{\frac{1}{2}}. \tag{54}$$

Eq. (53) shows that the dielectric constant of a crystal free to move is related to the dielectric constant of a clamped crystal by

$$K_F=K/(1-k^2). \tag{55}$$

For dynamic conditions, Eq. (52) can be written

$$\frac{\partial Q}{\partial y}=\frac{l_w K}{4\pi}\left[\frac{E}{l_t}-\frac{4\pi d_{14}}{2s_{22}'K}\frac{\partial \xi}{\partial y}\right]. \tag{56}$$

Eqs. (40), (50) and (56) give sufficient data to solve the problem. They are collected together in one equation for reference.

$$\rho l_w l_t dy\frac{\partial^2 \xi}{\partial t^2}=F_1-F_2=\frac{-\partial F}{\partial y}dy, \quad \text{(A)}$$

$$-F=\frac{l_t l_w}{s_{22}'}\left[\frac{\partial \xi}{\partial y}+\frac{2\pi d_{14}}{l_w K}\frac{\partial Q}{\partial y}\right], \quad \text{(B)} \tag{57}$$

$$\frac{\partial Q}{\partial y}=\frac{l_w K}{4\pi}\left[\frac{E}{l_t}-\frac{4\pi d_{14}}{2Ks_{22}'}\frac{\partial \xi}{\partial y}\right]. \quad \text{(C)}$$

Differentiating (57B) with respect to y and substituting the result in Eq. (57A) we have

$$\rho\frac{\partial^2 \xi}{\partial y^2}=\frac{1}{s_{22}'}\left[\frac{\partial^2 \xi}{\partial y^2}+\frac{4\pi d_{14}}{2l_w K}\frac{\partial^2 Q}{\partial y^2}\right]. \tag{58}$$

For a plated crystal E the applied potential is constant along the crystal and hence is not a function of y. Differentiating (57C) by y we have

$$\frac{\partial^2 Q}{\partial y^2}=-\frac{d_{14}}{2s_{22}'}l_w\frac{\partial^2 \xi}{\partial y^2},$$

which simplifies (58) to the form

$$\rho\frac{\partial^2 \xi}{\partial t^2}=\frac{(1-k^2)}{s_{22}'}\frac{\partial^2 \xi}{\partial y^2}. \tag{59}$$

For simple harmonic motion, (59) is satisfied by the relation

$$\xi=A\cos\frac{\omega}{v'}y+B\sin\frac{\omega}{v'}y \tag{60}$$

where $v'=((1-k^2)/\rho s_{22}')^{\frac{1}{2}}$.

When $y=0$, $A=\xi_1$ the displacement at the point $y=0$. Differentiating ξ by y we have

$$\frac{\partial \xi}{\partial y}=-A\frac{\omega}{v'}\sin\frac{\omega}{v'}y+B\frac{\omega}{v'}\cos\frac{\omega}{v'}y.$$

Hence

$$B=\left(\frac{\partial \xi}{\partial y}\right)_1\frac{v'}{\omega}.$$

Solving (57B) and (57C) simultaneously for $\partial\xi/\partial y$ we find

$$-\frac{\partial\xi}{\partial y}=\frac{s_{22}'}{l_t l_w(1-k^2)}\left[F+\frac{d_{14}}{2s_{22}'}l_w E\right]. \quad (61)$$

Inserting these values of A and B in (60), we have

$$\xi=\xi_1\cos\frac{\omega}{v'}y-\frac{[F_1+(d_{14}/2s_{22}')l_w E]}{l_t l_w\omega(\rho(1-k^2)/s_{22}')^{\frac{1}{2}}}\sin\frac{\omega}{v'}y. \quad (62)$$

From (60) and (61) we have also the relation

$$\left(F+\frac{d_{14}}{2s_{22}'}l_w E\right)=\left(F_1+\frac{d_{14}}{2s_{22}'}l_w E\right)\cos\frac{\omega}{v'}+\xi_1\omega l_t l_w\left(\frac{\rho(1-k^2)}{s_{22}'}\right)^{\frac{1}{2}}\sin\frac{\omega y}{v'}. \quad (63)$$

A third relation can be obtained by integrating (57C) directly. Since E is independent of y, we have

$$Q=\frac{l_w l_y K E}{4\pi l_t}-\frac{d_{14}l_w}{2s_{22}'}(\xi_2-\xi_1), \quad (64)$$

where l_y is the length of the crystal along the direction of vibration, and ξ_2 and ξ_1 are the displacements on the two ends of the crystal.

These three equations, which will determine the impedance of a crystal for any mechanical terminating conditions, are gathered together as Eq. (65). In this equation, however, we have replaced the particle displacement by the particle velocity and the charge by the current which can be done for simple harmonic motion by dividing the displacement and charge by $j\omega$. We have then

$$\dot{\xi}=\dot{\xi}_1\cos\frac{\omega}{v'}y-\frac{j[F_1+(d_{14}/2s_{22}')l_w E]}{l_t l_w(\rho(1-k^2)/s_{22}')^{\frac{1}{2}}}\sin\frac{\omega}{v'}y \quad (A)$$

$$\left[F+\frac{d_{14}}{2s_{22}'}l_w E\right]=\left(F_1+\frac{d_{14}l_w}{2s_{22}'}E\right)\cos\frac{\omega}{v'}y-j\dot{\xi}_1 l_t l_w\left(\frac{\rho(1-k^2)}{s_{22}'}\right)^{\frac{1}{2}}\sin\frac{\omega y}{v'} \quad (B) \quad (65)$$

$$E(j\omega C_0)=i+\frac{d_{14}l_w}{2s_{22}'}(\dot{\xi}_2-\dot{\xi}_1), \quad (C)$$

where C_0 is the total capacity of the crystal clamped, i.e., $(Kl_w l_y/4\pi l_t)$. These equations are of interest for supersonic phenomena as well as to determine the impedance of a crystal element. An example is considered later.

For a freely vibrating crystal, the forces on the two ends are zero. Introducing these values in (65A) and (B) we find

$$\dot{\xi}_1=-\dot{\xi}_2=\frac{jd_{14}E(1-\cos\omega l_y/v')}{2l_t(\rho s_{22}'(1-k^2))^{\frac{1}{2}}\sin\omega l_y/v'}=\frac{jd_{14}E}{2l_t(\rho s_{22}'(1-k^2))^{\frac{1}{2}}}\tan\frac{\omega l_y}{2v'}. \quad (66)$$

Inserting these values in Eq. (65C) and solving for the ratio of E/i, we find the impedance of the crystal to be

$$Z_C=\frac{E}{i}=\frac{-j}{\omega C_0\left[1+\dfrac{[k^2/(1-k^2)]}{\omega l_y/2v'}\tan\dfrac{\omega l_y}{2v'}\right]}. \quad (67)$$

The first electrical resonance will occur when

$$\tan\frac{\omega l_y}{2v'}=\infty \quad \text{or} \quad f_R=\frac{1}{2l_y}\left(\frac{(1-k^2)}{\rho s_{22}'}\right)^{\frac{1}{2}}. \quad (68)$$

This is lower than the natural mechanical resonance of the crystal by the factor $(1-k^2)^{\frac{1}{2}}$ in agreement with the experimental data of Fig. 1. When the denominator is zero, the impedance will be infinite and an antiresonance will occur. The frequency of antiresonance is given by the formula

$$\frac{\omega l_y}{2v'}\cot\frac{\omega l_y}{2v'}=-\frac{k^2}{1-k^2}. \quad (69)$$

Fig. 7 shows the calculated resonant and antiresonant frequencies plotted as a ratio to the natural mechanical resonant frequency for a given electromechanical coupling. It will be noted that both the resonant and antiresonant frequencies are lower than the natural mechanical

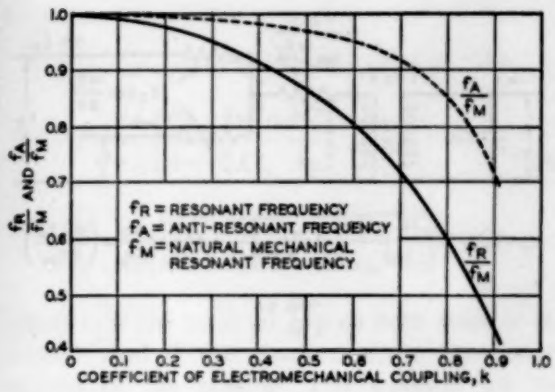


FIG. 7. Theoretical curves showing, respectively, the ratio between the electrical resonant and the natural mechanical resonant frequencies and the ratio between the electrical antiresonant frequency and the natural mechanical resonant frequency plotted as a function of the electromechanical coupling.

resonance of the crystal by considerable amounts when the coupling is large. Comparing the ratio of antiresonance to resonance for the measured curves of Fig. 1 and the theoretical curves of Fig. 7, the coupling can be evaluated for a Rochelle salt crystal as a function of temperature. This curve is plotted on Fig. 2. Knowing the coupling the natural mechanical resonant frequency can be calculated from Eq. (68). A calculation of this quantity for the crystal of Fig. 1 is shown by the dot-dash line labeled *D* of Fig. 1. This is slightly lower than that measured in an air-gap holder, but agrees in shape very well with that curve. The slight lowering is probably due to the weight of the plating used to drive the crystal. Hence the theory presented here agrees very well with the measured results.

In the introduction it was shown experimentally that the measured resonance and antiresonance frequency of a crystal approached the natural mechanical resonance frequency when the amount of plating was small. This can be proved theoretically by the use of Eqs. (65). We consider first the case of a crystal cemented on its two ends to two identical bars which it drives in longitudinal motion. Assuming that the cement is very thin and stiff, so that its mechanical impedance can be neglected, the crystal will drive two bars free on one end whose mechanical impedances are known to be[12]

$$Z_M = jZ_0 \tan \omega l_1/v_1, \quad \text{where} \quad Z_0 = l_{t_1} l_{w_1} \rho_1 v_1. \tag{70}$$

In this equation l_{t_1} is the thickness of the bars, l_{w_1} their width, and l_1 their length (measured in centimeters), ρ_1 the density and v_1 the velocity of propagation in the attached bars. Since this is the mechanical impedance that the crystal is terminated in on each side, we have the termination conditions

$$\frac{F_2}{\xi_2} = \frac{-F_1}{\xi_1} = jZ_0 \tan \frac{\omega l_1}{v_1}. \tag{71}$$

The negative sign arises since the ratio of force to velocity for the assumed directions are opposite for the two ends.

Inserting these ratios in Eqs. (65A) and (65B) and solving for the electrical impedance from (65C) we find

$$Z = \frac{E}{i} = \frac{-j}{\omega C_0 \left[1 + \dfrac{[k^2/(1-k^2)]}{\omega l_y/2v'} \left(\dfrac{\tan \omega l_y/2v'}{1 - Z_0/Z_0' \tan \omega l_y/2v' \tan \omega l_1/v_1} \right) \right]}, \tag{72}$$

where Z_0' is the mechanical impedance of the crystal, i.e., $l_t l_w \rho v'$. The resonant frequencies of the combination are given by

$$\tan \frac{\omega l_y}{2v'} \tan \frac{\omega l_1}{v_1} = \frac{Z_0'}{Z_0} = \frac{l_t l_w \rho v'}{l_{t_1} l_{w_1} \rho_1 v_1}, \tag{73}$$

while the antiresonant frequencies are given by the equation

$$\cot \frac{\omega l_y}{2v'} - \frac{Z_0}{Z_0'} \tan \frac{\omega l_1}{v_1} = -\left[\frac{k^2/(1-k^2)}{\omega l_y/2v'} \right]. \tag{74}$$

TABLE IV.

PERCENT PLATING	f_R/f_M	f_A/f_M	r CALCULATED
100	0.9951	0.99910	124.5
80	.99523	.99978	109.0
60	.99562	1.0000	113.3
40	.99663	1.0000	143.0
20	.99810	1.0000	262.0
0	1.0000	1.0000	∞

[12] This is easily proved from (65B) by letting d_{14} and hence $k=0$ and setting $F_2=0$. The result is $F_1/\xi_1 = jZ_0 \tan \omega l_1/v_1$ where $Z_0 = l_t l_w (\rho/s_{22}')^{\frac{1}{2}}$, $v_1 = (1/\rho s_{22}')^{\frac{1}{2}}$.

The case of interest is the case where the crystal is only partially plated in such a manner that a certain percentage of the length on each end is left unplated. For this case the cross-sectional area is constant down the whole length and the ratio

$$Z_0'/Z_0=(1-k^2)^{\frac{1}{2}}. \tag{75}$$

For X cut quartz crystals for example with their lengths along the Y or mechanical axis, the value of k has been found to be 0.099. A calculation was made of the ratio of the resonant frequency to the natural mechanical resonant frequency for various percentages of the crystal plated and the results are shown in Table IV.

This table shows also the ratio of the antiresonant frequency to the natural mechanical resonant frequency of the crystal while the third column gives the ratios of capacitances C_2 to C_1 in the equivalent electrical circuit of the crystal shown on Fig. 8. It is obvious that the crystal is most efficiently driven when it is about 70 percent plated, and this deduction has been verified experimentally.[13] The same conclusions hold for any other type of piezoelectric crystal. It will be noted that when the percent of plating becomes small both the resonant and antiresonant frequencies approach the natural mechanical resonant frequency f_m. The antiresonant frequency agrees with f_m when the amount of plating is 50 percent or less.

Another question of importance which can easily be solved by this method is the effect of an air gap between the crystal and the driving plates on the frequency of the crystal. The method of deriving the equations is the same as above, except that we have to calculate what part of the total potential is applied across the crystal and what part in the air gap. Eqs. (57)

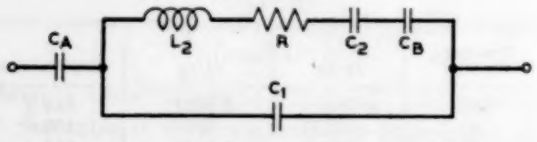


FIG. 8. An electrical circuit representing the electrical impedance of a piezoelectric crystal.

[13] This result was first shown experimentally by Mr. R. A. Sykes.

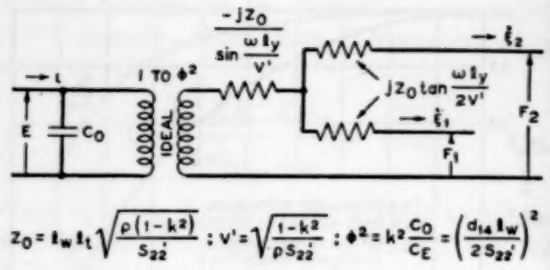


FIG. 9A

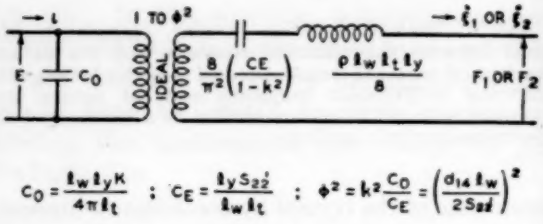


FIG. 9B

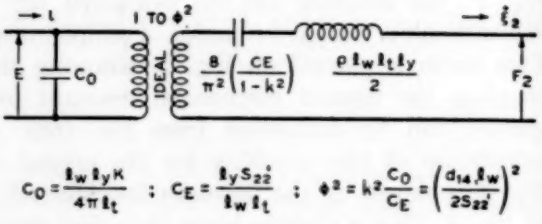


FIG. 9C

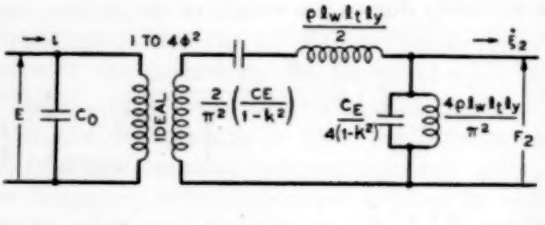


FIG. 9D

FIG. 9. An electromechanical representation of a piezoelectric crystal. (A) An exact representation, (B) a lumped constant representation for a symmetrical crystal, (C) a lumped constant representation for a crystal clamped on one end, (D) a lumped constant representation for a crystal free on one end and driving a load on the other.

become

$$\rho l_w l_t dy(\partial^2\xi/\partial t^2)=(-\partial F/\partial y) \quad \text{(A)}$$

$$-F=\frac{l_t l_w}{s_{22}'}\left[\frac{\partial\xi}{\partial y}+\frac{4\pi d_{14}}{2Kl_w}\frac{\partial Q}{\partial y}\right], \quad \text{(B)} \qquad (76)$$

$$\frac{\partial Q}{\partial y}=\frac{l_w l_t K}{4\pi(Kl_A+l_t)}\left[\frac{E}{l_t}-\frac{4\pi d_{14}}{2Ks_{22}'}\frac{\partial\xi}{\partial y}\right], \quad \text{(C)}$$

where l_A is the total air gap on both sides of the crystal, and Q the charge on the elementary section of the plating. The solution for the impedance of the crystal becomes

$$Z_c=\frac{-j}{\omega C_0'\left[1+\dfrac{[k'^2/(1-k'^2)]\tan\omega l_y/2v''}{\omega l_y/2v''}\right]}, \qquad (77)$$

where $$C_0'=\frac{Kl_w l_y}{4\pi(Kl_A+l_t)};$$

$$k'^2=\left(\frac{l_t}{Kl_A+l_t}\right)\frac{4\pi(d_{14}/2)^2}{Ks_{22}'};$$

$$v''=((1-k'^2)/\rho s_{22}')^{\frac{1}{2}}.$$

Hence the effect of an air gap is to change the static capacitance and the square of the electromechanical coupling by the factor $l_t/(Kl_A+l_t)$. In the limiting case when l_A becomes large, the coupling k' approaches zero and hence the measured resonance of the crystal will approach the natural mechanical resonance of the crystal. It is, therefore, valid to measure the elastic constants as described above in an air-gap holder.

An equivalent electrical circuit which will represent the impedance of Eq. (77) in the neighborhood of the resonant and antiresonant frequencies is shown on Fig. 8. These elements have the values

$$C_1=\frac{Kl_w l_y}{4\pi l_t};\quad C_2=\frac{8k^2C_1}{\pi^2(1-k^2)};$$

$$C_A=\frac{l_w l_y}{4\pi l_A};\quad C_B=\frac{8(C_1+C_A)}{\pi^2-8};\quad L_2=\frac{s_{22}'\rho l_y^2}{8k^2C_1}. \qquad (78)$$

The resistance shown represents the dissipation put in by internal losses, clamping losses, and all other sources.

An exact electromechanical representation[14] of a plated piezoelectric crystal is shown on Fig. 9(A). If one solves for the equations of this network, they reduce to those given by Eqs. (65). If we only wish to use the representation near the resonant frequency, lumped network representations can be used. Fig. 9(B) shows a representation for a symmetrical crystal driving equal loads on each end. Fig. 9(C) shows a representation for a crystal clamped on one end and driving a load on the other end, while Fig. 9(D) shows a representation for a crystal free on one end and driving a load on the other (inertia drive).

[14] An approximate representation is discussed in a previous paper "An Electro-mechanical Representation of a Piezo-Electric Crystal used as a Transducer," W. P. Mason, Proc. I. R. E. **23**, 1252 (1935).

LETTERS TO THE EDITOR

Prompt publication of brief reports of important discoveries in physics may be secured by addressing them to this department. Closing dates for this department are, for the first issue of the month, the eighteenth of the preceding month, for the second issue, the third of the month. Because of the late closing dates for the section no proof can be shown to authors. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents.

Communications should not in general exceed 600 words in length.

Binding Energies of Light Nuclei

In the light nuclei beyond He^4, the Hartree approximation is known to give energies far from the experimental values. In the absence, however, of more extended work upon most of these nuclei, the Hartree calculations of Feenberg, Wigner, and Phillips[1] have been the best available. We have now completed extensive calculations upon all the nuclei through Li^7. The method employed has been outlined in previous papers.[2] The purpose of this note is to present for different nuclei a general comparison of our results,[3] which is of interest chiefly because of the uniformity of procedure applied throughout.

The symmetric Hamiltonian has been used, in which the interaction between all pairs of nuclear particles is

$$V_{ij} = -A \exp(-r_{ij}^2/a^2)(w + mP_{ij} + bQ_{ij} + hP_{ij}Q_{ij}). \quad (1)$$

We have employed the following values for the nuclear parameters:

$A = 35.60$ Mev; $a = 2.25 \times 10^{-13}$ cm;
$w = -2/15$; $m = 14/15$; $b = 7/15$; $h = -4/15$.

This set of values was found to lead in general to higher binding than other proposed sets compatible with the usual inequalities.[2]

Convergence limits for H^3 and He^4 have been carefully estimated by means of refined variational calculations. For He^5, He^6, Li^6, Li^7, upper bounds to the second-order perturbation energy have been obtained. The relation of these upper bounds to the ultimate theoretical energies has been investigated in various ways.

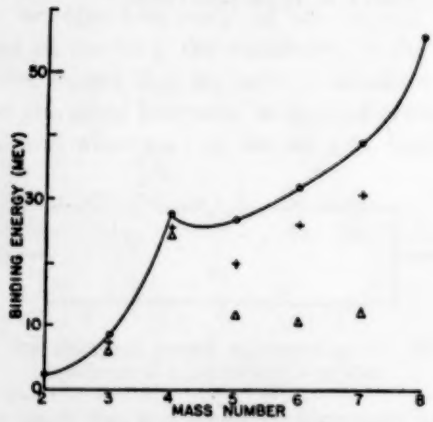


FIG. 1. Binding energies of light nuclei. Triangles—Hartree approximation; crosses—maximum binding in higher approximations; circles—experimental values. Only the most stable nucleus for each mass number is plotted.

TABLE I. *Cases in which the intervals between levels are altered by higher approximation.*

INTERVAL	HARTREE APPROXIMATION	HIGHER APPROXIMATION
He^5: $^2S-^2P$	5 Mev	~1 Mev
Li^6: $^3D-^3S$	2 Mev	6 Mev
$^1S(He^6)-^3S(Li^6)$	3.2 Mev	2.9 Mev

Figure 1 summarizes the results. Beyond He^4 the Hartree model gives only about one-third of the experimental energy, the higher approximations contributing about another third.

The following remarks may now be made about the validity of the Hartree approximation. (1) Higher approximations contribute as much as, or more than, the Hartree method. (2) The scalloped structure, so manifest in the experimental binding-energy curve, and incipiently present in the Hartree approximation, is even more pronounced when higher approximations are included. (3) Although the Hartree model appears to give correctly the *order* of low lying levels, we find in two specific cases that the intervals between these levels are altered radically upon higher approximation. Numerical results for these instances are given in Table I. On the other hand the difference $^1S(He^6)-^3S(Li^6)$ is *not* appreciably altered in higher approximation, which is to be attributed to the fact that these two levels possess the same orbital symmetry.

In addition, one may draw the conclusion from the present work (as well as from other recent developments) that the symmetric Hamiltonian with interaction (1) or any reasonable modification of it is not entirely adequate to explain the binding energies of nuclei immediately beyond He^4.

W. A. TYRRELL, JR.
K. G. CARROLL
H. MARGENAU

Sloane Physics Laboratory,
Yale University,
New Haven, Connecticut,
March 30, 1939.

[1] E. Feenberg and E. Wigner, Phys. Rev. **51**, 95 (1937); E. Feenberg and M. Phillips, Phys. Rev. **51**, 597 (1937).
[2] H. Margenau and D. T. Warren, Phys. Rev. **52**, 790 (1937); H. Margenau and W. A. Tyrrell, Jr., Phys. Rev. **54**, 422 (1938); H. Margenau and K. G. Carroll, Phys. Rev. **54**, 705 (1938).
[3] Cf. Abstracts 31 and 32, New York Meeting, American Physical Society, February 23–25, 1939, and Abstract No. 94, Washington Meeting, American Physical Society, April 27–29, 1939.

Energy Production in Red Giants

It was indicated by one of us[1] that in order to explain the energy production in red giants it is necessary to accept the presence in their interior of considerable amounts of *light elements* reacting with hydrogen at comparatively low temperatures and densities. In this case one should expect that the distribution of stars in the giant region of the Russell diagram should show a *"lateral" structure;* the stars with different nuclear reactions are located in different bands parallel to the main sequence. We have investigated theoretical location of such bands and compared it with the observational evidence.

We assume that the energy production in the main sequence is entirely due to the N−C cycle[2] and accept the conventional point-source equilibrium equations for the stellar interior, with the total pressure given by the sum of gas and radiative pressures. For constant mass and hydrogen content but variable radius ($R=\alpha R_0$) we obtain for the homology transformations[1] for temperature and luminosity the expressions:

$$T=\alpha^{-1}T_0, \tag{1a}$$

$$L=\alpha^{-1.25}L_0, \tag{1b}$$

where index "zero" corresponds to a star of the main sequence. This particular transformation holds whatever the relative contributions of gas and radiative pressure may be.

For the relative energy production we have, on the other hand:

$$\frac{L}{L_0}=\left(\frac{T}{T_0}\frac{Z_0^2m_0}{Z^2m}\right)^n\frac{\epsilon\rho\Gamma}{\epsilon_0\rho_0\Gamma_0}, \tag{2}$$

where ϵ is the energy released per penetration, Γ the capture probability, and n the exponent for the temperature dependence in the region under discussion. Substituting (1a) and (1b) into (2) we obtain:

$$\alpha=\left(\frac{Z_0^2m_0}{Z^2m}\right)^{n/(n+1.75)}\left(\frac{\epsilon\Gamma}{\epsilon_0\Gamma_0}\right)^{1/(n+1.75)}. \tag{3}$$

The reactions of light elements which come into consideration are listed in Table I, together with the values of ϵ and Γ.[3] In the last column are given the calculated values of lg α which define the positions of the bands parallel to the main sequence in the radius-luminosity diagram, along which the stars with different nuclear reactions should be located.

For comparison with the observational material we choose the pulsating stars. As is well known, these form a one-dimensional sequence, and all of their characteristics

TABLE I. *Energy release and capture probability for some of the reactions between light elements.*

REACTION	ϵ(×10⁶ EV)	Γ(EV)	lg α
$_1D^2+{}_1H^1 \rightarrow {}_2He^3+h\nu$	5.9	7	1.61
$_1T^3+{}_1H^1 \rightarrow {}_2He^4+h\nu$	21.3	10	1.60
$_3Li^6+{}_1H^1 \rightarrow {}_2He^4+{}_2He^3$	4.1	$5\cdot10^3$	0.94
$_3Li^7+{}_1H^1 \rightarrow {}_2He^4$	18.6	$4\cdot10^4$	0.91
$_4Be^9+{}_1H^1 \rightarrow {}_3Li^6+{}_2He^4$	2.4	10^6	0.73
$_5B^{10}+{}_1H^1 \rightarrow {}_6C^{11}+{}_2He^4$	9.2	10	0.22
$_5B^{11}+{}_1H^1 \rightarrow 3{}_2He^4$	9.4	10^6	0.53
$_7N^{14}+{}_1H^1 \rightarrow {}_8O^{15}+h\nu$	7.8	5	0.00

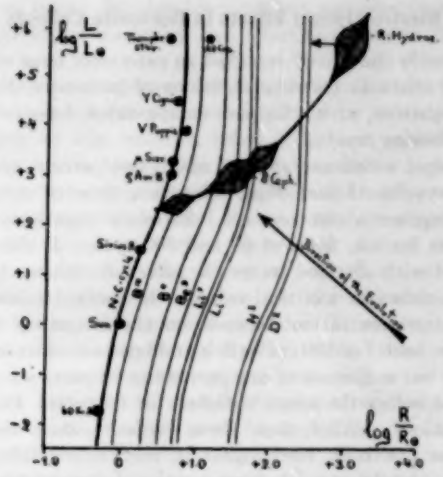


FIG. 1. Pulsating variables and different nuclear reactions.

are definite functions of the vibration period (we do not attempt here to explain the cause of pulsations, but choose pulsating stars only because they can be arranged in the sequence as the function of an exactly observable parameter related to their mass). In Fig. 1 the region of pulsating stars is shown by the shaded area; its width at each point is proportional to the number of stars observed.[3] We see that there are definite concentrations of stars near the places where different nuclear reactions of light elements, calculated entirely without adjustable parameters, become of importance. The short period group of cluster variables contains the stars in which the transformation of B^{10} probably plays the important role. The second group includes the Cepheid variables, the number of which show a broad and irregular distribution as a function of the period; this is probably due to the overlapping of three separate groups corresponding to the reactions of Li, Be and B^{11}. Finally the group of long period or "mira-variables" should be interpreted as due to deuterium (and possible tritium) reaction. It can be seen that the calculated radii for the stars of this group are about five times smaller than the observed radii. The same discrepancy, however, is also well known in the calculation of their radii from the period-density relation ($P\sim\rho^{-\frac{1}{2}}$). Its explanation can be seen in the fact that, according to calculations of Chandrasekhar,[4] these stars show *abnormally large central condensations;* 90 percent of their mass is concentrated within ten percent of the radius (instead of 50 percent of the radius as in normal stars). Thus, the large radii observed actually correspond to rarefied atmospheres surrounding the main body and cannot be used for the calculation of central temperatures and periods of vibration.

G. GAMOW
E. TELLER

The George Washington University,
Washington, D. C.,
March 20, 1939.

[1] G. Gamow, Phys. Rev. 55, 718 (1939).
[2] H. Bethe, Phys. Rev. 55, 434 (1939).
[3] H. N. Russell, R. Dugan and J. Stewart, *Astronomy* (Ginn and Company, 1927), p. 759.
[4] S. Chandrasekhar, *Stellar Structure* (Chicago University Press, 1939) p. 318.

Electro-Optical Effects in Bentonite Colloids

Recently the writer[1] reported an extremely large electro-optical effect in colloidal solutions of bentonite. Further investigations, with a Babinet compensator, have revealed the following results.

A sol of a concentration of about one percent (particle size between 15 and 30$\mu\mu$, relaxation time of streaming birefringence about one sec.) becomes negatively birefringent for a.c. fields of 60 and 500 cycles. If this sol is diluted with distilled water the effect diminishes rapidly and vanishes for a critical value of the concentration. This critical concentration depends on the frequency of the electric field. For 500 cycles it is at 38 percent of its original value, but a dilution of one part sol in 80 parts water was needed before the effect vanished for 60 cycles. For concentrations smaller than these critical values the sols become positively birefringent in the electric field. This positive birefringence is largest when the concentration is about one-half of the above critical concentrations.

These observations confirm therefore the results of F. J. Norton[2] who reports a change from negative to positive birefringence when the concentration is kept constant and the frequency is increased. Our results show that the critical frequency for which the birefringence disappears increases with the concentration of the sol. It may be as small as 60 cycles for dilute solutions and it can be raised to high frequencies in concentrated sols, but so far we have not been able to get a reversal for 10^6 cycles, for which the birefringence is positive even for the very concentrated sols.

These results make it most unlikely that the phenomenon is analogous to the reversal of the Kerr effect in rosin and octyl alcohol. For these polar liquids the critical frequency increases with decreasing viscosity while in the clay colloids the most dilute and least viscous solutions have the lowest critical frequency. The tansition from negative to positive birefringence is not a relaxation phenomenon.

The relaxation time of the birefringence has been investigated. As reported previously the double refraction does not follow the alternations of the field when concentrated solutions are used. However, when the concentration is decreased, one reaches a dilution where the bentonite "Kerr" cell acts as an electro-optical light shutter.

In observations with a Babinet compensator the interference lines gradually become blurred when the electric field is increased and experiments with a photo-cell amplifier connected to a cathode-ray oscillograph give a clear record of the fluctuations of the light intensity. We find that the birefringence consists in general of two parts, one which is constant and one which vibrates with twice the frequency of the electric field. The constant and alternating parts of the birefringence have always the same sign and they vanish both at the same concentration. For concentrated solutions the alternating part is negligible, for dilute solutions the constant part disappears. The concentrations for which the alternating part becomes noticeable are about 0.8 percent for 60 cycles and 0.5 percent for 500 cycles, i.e., the lower the concentration the higher is the critical frequency at which the cell begins to act as a stroboscope. The disappearance of the alternating birefringence follows all the laws of a relaxation effect, it depends on temperature and viscosity, but there is no connection between this effect and the reversal of birefringence.

The discovery of the alternating part of the birefringence makes it possible to use bentonite in light shutters and stroboscopes. They can be operated on small voltages and be used for wide light beams. Even in the dilute solutions which are almost as transparent as pure water the "Kerr" effect is still many thousand times larger than in nitrobenzene. From a theoretical point of view, however, these new observations complicate the problem of the origin of the electro-optical effects in colloids.

HANS MUELLER

Massachusetts Institute of Technology,
Cambridge, Massachusetts,
March 18, 1939.

[1] H. Mueller, Phys. Rev. **55**, 508 (1939).
[2] F. J. Norton, Phys. Rev. **55**, 668 (1939). A larger part of our results were obtained before Dr. Norton informed me of his work.

The Neutron-Proton Scattering Cross Section

The cross section for the scattering of thermal neutrons by protons in paraffin wax has been determined rather accurately (for literature see reference 1). Taking into account the chemical binding forces acting on the protons in the paraffin, one can from these measurements calculate the scattering cross section of free protons relative to slow neutrons; in this way, a value of 14×10^{-24} cm^2 has been obtained.[1, 2] This cross section has also been measured directly by studying the activity produced in different elements when activated through a shield of cadmium.[3] The mean free path in paraffin, of the "resonance" neutrons of these elements, was found to be 1.1 cm, which corresponds to a value about 11.5×10^{-24} cm^2 for the scattering cross section of the proton. The poor agreement between the calculated and the experimental value and the importance of an accurate knowledge of the neutron-proton cross section for a number of nuclear problems made a new and more accurate determination seem highly desirable. In the following a short report of such measurements is given, a detailed account of which is to appear soon in the *Proceedings of the Royal Danish Academy*, Copenhagen.

The measurements were carried out with water as a proton scatterer and with both silver and iodine as detector and filter materials. The source (300–600 mg radium plus beryllium) was placed three cm below the upper surface of a cube of paraffin wax of ten cm sides. The detector, in a pocket of Cd-sheet (0.4 g/cm^2), was placed 5 to 14 cm above the upper surface of the paraffin cube, the scatterer midway between source and detector, and the filter immediately below the scatterer. The areas of detector, scatterer and filter were 10×10 cm^2. Runs (of equal duration) were then made (1) with only the empty water trough between paraffin and detector, (2) with the filter, (3) with the water scatterer, and (4) with filter and water scatterer. If the recorded numbers of particles are called

N_1, N_2, N_3, N_4, respectively, the ratio $(N_3-N_4)/(N_1-N_2)$ gives the transmission, in the water layer, of those resonance neutrons which have suffered no energy change of more than the order of the resonance width.

Results obtained with different detectors (silver and iodine), different water layers (5, 10 and 15 mm) and different distances of the detector from the paraffin block (5 and 14 cm) agreed within the statistical errors, and the weighted mean for the mean free path of the resonance neutrons in water was found to be 0.91 cm, which corresponds to a proton cross section of 14.8×10^{-24} cm² if the cross section of oxygen is taken as 3.3×10^{-24} cm².[4] The standard error of these figures is about five percent.

Using essentially the same method, Cohen, Goldsmith and Mitchell found a considerably larger value (20×10^{-24} cm²),[5] with rhodium as detector. It should be remembered, however, that the resonance energy of rhodium is about one ev only, which is not much larger than the quantum energy of some of the proton oscillation frequencies, and that one must therefore expect the rhodium resonance neutrons to be scattered more strongly than slightly faster ones. Indeed the silver resonance neutrons were found, in the present experiment, to be scattered practically in the same way as the considerably faster resonance neutrons of iodine.

In conclusion I wish to express my sincere thanks to Professor N. Bohr for the opportunity to work at the Institute of Theoretical Physics, and to Dr. O. R. Frisch for valuable advice.

LENNART SIMONS

Institute of Theoretical Physics,
University, Copenhagen,
March 10, 1939.

[1] H. A. Bethe, Rev. Mod. Phys. **9**, 69 (1937).
[2] Arley, Kgl. Danske Vidensk. Selsk. Mat.-fys. Med. **16**, 1 (1938).
[3] E. Amaldi and E. Fermi, Phys. Rev. **50**, 899 (1936).
[4] Dunning, Pegram and Mitchell, Phys. Rev. **48**, 277 (1935).
[5] Cohen, Goldsmith and Schwinger, Phys. Rev. **55**, 106 (1939).

The Cross Sections of Metallic Uranium for Slow Neutrons

Whitaker and Beyer[1] have shown that slow neutron cross sections are not always additive. Their data indicate that the cross sections obtained by measurements on compounds may differ greatly from those obtained by measurements made on the elements. Because of this variation in results obtained by direct and indirect measurements and because of the current interest in uranium, it seemed worth while to measure the cross section of uranium for slow neutrons with the use of the metal instead of the compounds used heretofore. The availability of one pound of uranium metal enabled us to do this. It seemed especially desirable to make measurements that will enable one to determine not only the total cross section but also the capture cross section and the scattering cross section. With this aim in view the experiments described below were done.

The uranium metal was packed in an aluminum box 8×8 cm and 0.8 cm thick. The sides of the box were made of 0.16-cm aluminum sheets. This box held 6.85 g/cm² of uranium metal. Another identical box was used in making corrections for scattering and adsorption due to the box itself.

The experiments consisted of measuring the transmission of the uranium for the slow neutrons by detecting the intensity of the neutron beam with and without the uranium in the beam. The sample was far enough away from the detector so that no appreciable number of scattered neutrons were counted. These measurements give the total thermal neutron cross section for capture and for scattering. By repeating these measurements with the uranium sample very close to the detector one counts a fraction of the scattered neutrons. This fraction may be determined experimentally by using any material whose capture cross section and scattering cross section are both known. The capture cross section of carbon is believed to be so small as to be negligible. The capture cross section and the scattering cross section for iron have been determined and found to be 3.5 and 8.5×10^{-24} cm², respectively, by this method of change in solid angle.[2] Both iron and carbon have been used in these experiments to make the necessary corrections. These included a correction for the change in the path length of the scattered neutrons in the detector.

The total cross section for *this* uranium metal was found to be $23.1\pm0.5\times10^{-24}$ cm². The analysis for this metal does not show impurities which should influence the cross section appreciably. This measurement of total cross section agrees reasonably well with the recent determination for uranium oxide made by Goldsmith, Cohen, and Dunning.[3] I am informed by Professor Dunning that they are reporting a cross section of $20.0\pm2.0\times10^{-24}$ cm². Through the courtesy of Professor Dunning a check was made for the value of the metal with the same neutron beam as was used in measuring the oxide. The result obtained was $23.3\pm0.5\times10^{-24}$ cm². The experiment for the separate determination of the capture cross section and the scattering cross section in which a silver detector was placed as close to the sample being studied as was possible, gave a result for the capture cross section of 11 ± 3, and for the scattering cross section of $12\pm3\times10^{-24}$ cm².

While it is true that these last measurements are less direct than those for the total cross section, it is believed that the limits of error put on these values are conservative. The results are being checked by a scattering experiment of the type done by Goldhaber and Briggs.[4]

A grant-in-aid for research from Sigma Xi to one of us (EJM) is gratefully acknowledged.

MARTIN D. WHITAKER
CHARLES A. BARTON
WILLIAM C. BRIGHT

New York University,
University Heights,

EDGAR J. MURPHY

College of the City of New York,
New York, New York,
April 1, 1939.

[1] Whitaker and Beyer, Abstract No. 72, Washington Meeting of the American Physical Society, April 27–29, 1939.
[2] Whitaker, Beyer and Dunning, Phys. Rev. **54**, 771 (1938).
[3] Goldsmith, Cohen and Dunning, Abstract No. 69, Washington Meeting of the American Physical Society April 27–29, 1939.
[4] M. Goldhaber and G. H. Briggs Proc. Roy. Soc. **162**, 127 (1937).

Radioactive Tellurium: Further Production and Separation of Isomers

We have found seven chemically identified radioactive periods of tellurium as the result of deuteron and neutron bombardment of tellurium, fast neutron bombardment of iodine, and deuteron and proton bombardment of antimony. This work has made it possible to make the isotopic assignments for all of the activities and to show that there are three pairs of isomers. The "isomer separation" method of Segrè, Halford and Seaborg,[1] as already applied to tellurium,[2] was tried for each of these pairs. It is remarkable that all three isomeric pairs were chemically separable, the short period growing from the long-lived activity in each case.

Te^{127}: *90 days and 10 hours.*—Prolonged bombardment of iodine with the fast neutrons from Li+D produces in the tellurium precipitate not only the electron-emitting ten-hour activity previously reported[3] but also a new activity with a 90-day half-life. Since there is only one stable iodine the reaction must be I^{127} (n, p) Te^{127} and both periods must be due to isomers of Te^{127}. These activities are produced with much greater intensity by deuteron bombardment of tellurium; it is possible to separate the ten-hour period from tellurium activated in this manner long after the directly formed ten-hour period has completely decayed (in one case, 45 days after the bombardment). The 90-day activity is due to the higher of the two isomeric levels and decays, probably by a converted gamma-ray transition, to the lower level, which then decays with a ten-hour half-life by beta-emission to stable iodine.

Te^{131}: *1.2 days and 25 minutes.*—We have already shown[4] that an electron-emitting eight-day iodine grows from a tellurium isotope, necessarily either Te^{129} or Te^{131}, with both a 1.2-day and a "short" half-life; further work[5] has demonstrated that the period of the shorter-lived parent of the iodine is about 25 minutes. The 1.2-day and 25-minute activities, which are both directly produced by deuteron bombardment of tellurium, are isomeric and isomer separation experiments show that the 25-minute period grows from the 1.2-day activity; it is in fact possible to observe, by successive extractions of iodine, the growth of the eight-day iodine from the 25-minute tellurium activity obtained by extraction from its parent isomer. Bothe and Gentner[6] did not find a 25-minute activity (but did observe a 60-minute period, which aids in assigning another pair of isomers to Te^{129}) when exposing tellurium to gamma-rays; this activation could produce Te^{129} but not Te^{131}, so it appears that the 1.2-day and 25-minute isomers must belong to Te^{131}.

Te^{129}: *30 days and 70 minutes.*—Deuteron bombardment of tellurium also produces a tellurium period of approximately 30-days half-life. Isomer separations show that a 70-minute tellurium activity grows from this 30-day period. The 70-minute activity is also undoubtedly produced directly by deuteron bombardment, but the presence of so many other activities makes it difficult to observe; however, an electron emitting activity with half-life about one hour seems to be present in the deuteron activated samples. If our electron-emitting 70-minute activity is to be identified with the 60-minute period that is produced by gamma-rays,[6] the only remaining possibility for the isotopic assignment is Te^{129}. This Te^{129} decays into I^{129} which should be unstable; the period of this iodine must be either very short or very long, since iodine extractions from the active tellurium have not revealed any daughter activity.

Te^{121}: *120 days.*—The activation of antimony with eight-Mev deuterons or with four-Mev protons gives in both cases a new chemically identified tellurium with the half-life about 120 days. The only reasonable reactions which would lead to unstable tellurium isotopes are Sb^{121} $(d, 2n)Te^{121}$ and Sb^{121} $(p, n)Te^{121}$. Low energy electrons are emitted. It is probable that the electrons result from an internally converted gamma-ray following the decay by K-electron capture to stable Sb^{121}. (Of course the possibility exists that the electrons come from a converted gamma-ray arising from a transition between two isomeric states in Te^{121} or in a stable tellurium isotope.)

Complete details of this work will be published at a later date. It is a pleasure to thank the Research Corporation for its continued support.

G. T. SEABORG
J. J. LIVINGOOD*
J. W. KENNEDY

Radiation Laboratory,
Department of Physics,
Department of Chemistry,
University of California,
Berkeley, California,
March 31, 1939.

* Now at Harvard University.

[1] E. Segrè, R. S. Halford and G. T. Seaborg, Phys. Rev. **55**, 321 (1939).
[2] G. T. Seaborg and J. W. Kennedy, Phys. Rev. **55**, 410 (1939).
[3] J. J. Livingood and G. T. Seaborg, Phys. Rev. **53**, 1015 (1938).
[4] J. J. Livingood and G. T. Seaborg, Phys. Rev. **54**, 775 (1938).
[5] W. Bothe and W. Gentner, Naturwiss. **25**, 191 (1937).

Isotopic Constitution of Hafnium, Yttrium, Lutetium and Tantalum

Hafnium ions were obtained from a high frequency spark between nickel tubes packed with a mixture of lanthanum metal and hafnium oxide. Doubly charged ions were photographed as they gave greater photographic intensity due to the greater accelerating voltage possible. In addition to the five isotopes reported by Aston, at 180, 179, 178, 177, and 176, a new isotope was found at 174 on six photographs. Its intensity was estimated at 0.3 percent of the total. That the line was not due to strontium (87) or rubidium (87) was shown by the absence of the stronger isotopes of these elements at 86 and 85, respectively. The mass could also be compared with the neighboring isotopes with sufficient accuracy to rule out the possibility of it belonging to these lighter elements which have a much smaller packing fraction.

On three of the photographs a much fainter mass near 86 was observed. This is possibly another isotope of hafnium at 172, but further experiments will be necessary to rule out the possibility of its being strontium (86).

With yttrium ions from a mixture of yttrium oxide and lanthanum metal it could be shown that no isotope existed at mass 91 as intense as 0.05 percent of the strong mass at 89.

A sample of lutetium oxide was kindly supplied by

Professor B. S. Hopkins of the University of Illinois and electrodes were made by packing nickel tubes with a mixture of the oxide and lanthanum metal. A heavier isotope was suggested by the atomic weight and the packing fraction curve.[1] Doubly charged ions were photographed, and it was found that the sample contained traces of ytterbium which unfortunately gave a mass at 176. On the other hand, no intensity occurred at mass 177 as strong as one one-thousandth of the intensity of the strong lutetium mass at 175. In the meantime, a second isotope has been found at 176 by Mattauch and Lichtblau[2] which gives good agreement with the atomic weight, and provides a striking exception to the usual isotopic structure of odd-numbered elements.

At the suggestion of Professor Hönigschmid, tantalum was examined, as its atomic weight with a packing fraction of 1.5×10^{-4} would imply the existence of a lighter isotope. If this isotope were at 179, it should have an intensity of approximately 5 percent of the main mass at 181. A spark between metallic tantalum electrodes served as a source, and singly, doubly, and triply charged ions were photographed. No mass could be found at 179 as strong as 0.1 percent of the main isotope at 181. A faint mass of approximately 0.15 percent was observed at 182 and also with doubly charged ions at 91. This is provisionally attributed to a hydride.

ARTHUR J. DEMPSTER

Ryerson Physical Laboratory,
University of Chicago,
Chicago, Illinois,
March 31, 1939.

[1] A. J. Dempster, Phys. Rev. **53**, 873 (1938).
[2] J. Mattauch and H. Lichtblau, Zeits. f. Physik **111**, 514 (1939).

On the Scattering of Fast Neutrons of Different Energy

In our previous papers,[1] the absorption scattering cross sections of atoms for D+D and Li+D neutrons were studied and it was found that cross sections depend upon the energy of neutrons. The neutrons were observed at right angle to the incident deuteron. In the present experiment, measurements were performed at various angles φ to the incident deuteron. The energy of the neutrons was changed between about 2.1 and 2.8 Mev. The accelerating voltage was 300 kv and the apparatus used was the same as that in previous work; the neutrons were detected by an ionization chamber, filled with methane at 20 atmospheres pressure, which was situated 30 cm from the target. The intensity of neutrons not coming from the target was estimated by inserting a thick paraffin block between the target and the chamber.

In calculating the mean energy of neutrons corresponding to various values of φ, it was assumed that our deuteron beam contained 30 percent atomic and 70 percent molecular ions and also that the target was thick. On account of the uncertainty of these assumptions, the energy values may be in error by about 0.15 Mev, except the value for $\varphi=90°$. But the relative spacing of the energy values are much more accurate. In correcting the transmission for the finite size of chamber and the absorber, the scattering of neutrons is assumed to be isotropic with respect to the center of gravity. The corrections were four percent of the scattered neutrons for atoms except H and D, and 16 percent and 9 percent for H and D, respectively. As the correction in the case of H is large, the experiment was repeated at 40 cm target-chamber distance where the correction was 9 percent, and the same result was obtained.

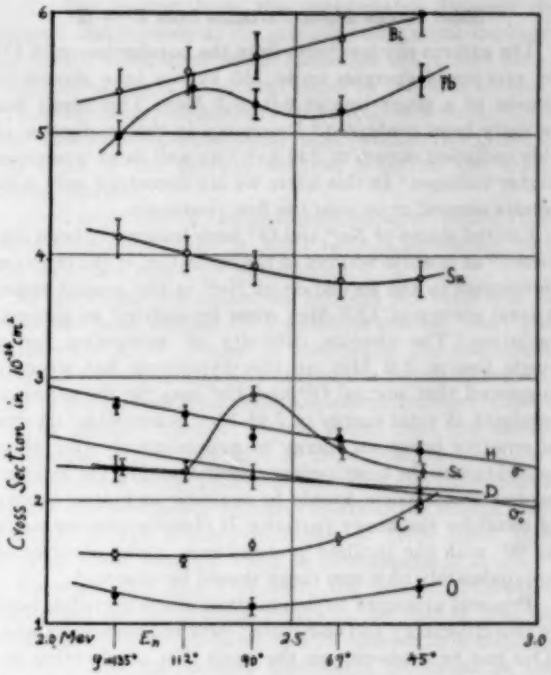


FIG. 1. Cross section for scattering of neutrons of various energies.

Thirty-two elements were examined and some of the results are shown in the figure. Paraffin and carbon were used for determining the cross sections of H and C, and D_2O, SiO_2 and Si were used to examine D, O_2 and Si. Complicated dependence of cross sections upon the energy of neutrons is observed in most cases. Si has a rather sharp resonance at 2.45 Mev. The cross section of H is larger than our previous value[1] (2.28 ± 0.09) and Ladenburg's[3] value (2.1 ± 0.2). The present value is not in agreement with the theoretical expression[3]

$$\sigma_0=\frac{4\pi\hbar^2}{M}\left(\frac{1}{4}\frac{1}{|E_1|+\frac{1}{2}E}+\frac{3}{4}\frac{1}{|E_0|+\frac{1}{2}E}\right)$$

but is in agreement with the following expression corrected for the final range of interaction of proton and neutron[4]

$$\sigma=\frac{4\pi\hbar^2}{M}\left(\frac{1}{4}\frac{1+\alpha_1 r_0}{|E_1|+\frac{1}{2}E}+\frac{3}{4}\frac{1+\alpha_0 r_0}{|E_0|+\frac{1}{2}E}\right)$$

when we assume $r_0=1.3\times10^{-13}$ cm.

The detailed report will soon appear in the *Proceedings of the Physico-Mathematical Society of Japan*.

HIROO AOKI

Physical Institute,
Osaka Imperial University,
Osaka, Japan,
March 5, 1939.

[1] Hiroo Aoki, Proc. Phys. Math. Soc. Jap. **21**, 40 (1939); Proc. Phys. Math. Soc. Jap. **21**, 75 (1939); Seishi Kipuchi and Hiroo Aoki, Phys. Rev. **55**, 108 (1939).
[3] R. Ladenburg and M. H. Kanner, Phys. Rev. **52**, 911 (1937).
[4] Cf. H. A. Bethe and R. F. Bacher, Rev. Mod. Phys. **8**, 117 (1936).
[4] Cf. H. A. Bethe and R. F. Bacher, Rev. Mod. Phys. **8**, 119 (1936).

Short Range Alpha-Particles from $F^{19}+H^1$

The gamma-ray spectrum from the bombardment of F^{19} by protons of energies up to 750 kv has been shown[1] to consist of a single line at 6.0 ± 0.2 Mev. This result has recently been confirmed.[2] Resonance in the production of this radiation occurs at 330 kv[3, 4] as well as at numerous higher voltages.[5] In this letter we are concerned only with results secured at or near the first resonance.

Excited states of Ne^{20} and O^{16} have frequently been discussed[6] as possible sources of the radiation. If the reaction terminates in the formation of Ne^{20} in the ground state, a total energy of 13.0 Mev must be emitted as gamma-radiation. The obvious difficulty of accounting for a single line at 6.0 Mev on this hypothesis has strongly suggested that normal O^{16} and He^4 may be the terminal products. A total energy of 7.95 Mev is available[7] on this alternative hence an energy of approximately two Mev, in addition to the kinetic energy contributed by the incident proton at resonance, should be available as kinetic energy of recoil for the heavy particles. If observations are made at 90° with the incident proton beam, alpha-particles of approximately nine mm range should be observed.

Previous attempts to detect these alpha-particles, both in this laboratory and elsewhere,[7] have not been successful. This can be explained on the basis that accelerating potentials of over 500 kv were employed and that for such high potentials the range of the protons scattered from the target is equal to or greater than that to be expected for the alpha-particles. The large ratio of scattered to disintegration particles almost entirely precludes the possibility of observing the latter under these conditions.

In the present investigation a small electrostatic generator was employed at and below the first gamma-ray resonance. A cloud chamber was used containing a mixture of helium and alcohol vapor having a stopping power approximately ten percent that of air under standard conditions. The disintegration products from a thick CaF_2 target inclined at 45° with the incident beam were admitted to the chamber through a thin mica or lacquer window (~3 mm equivalent range). Diaphragms limited the observed particles to those making an angle of $90°\pm8°$ with the incident beam.

In Fig. 1 is shown the number range curve of the tracks observed on 1800 cloud-chamber photographs with an accelerating potential of 350 kv and a proton current of one to five microamperes. The range scale was calibrated[8] by measuring under similar conditions the He^3 and He^4 groups from Li^6+H^1. The continuous distribution in Fig. 1 terminating at six mm is due to the protons scattered by the target. The mean range of the discrete group of particles is found to be 0.86 ± 0.05 cm. That this group is produced at a resonance voltage equal or very near to that of the first gamma-ray resonance was demonstrated by taking cloud-chamber pictures at 300 and 350 kv. On 200 pictures taken at 300 kv no tracks in this group were observed, while on 300 pictures taken immediately afterward at 350 kv 49 tracks of range near nine mm were observed. The fact that this group is much narrower than the Li^6+H^1 groups is also evidence for resonance. The observed track density is consistent with the assumption that

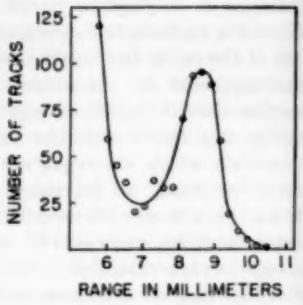


FIG. 1. The distribution in range of particles produced by the bombardment of F^{19} with 350-kv protons. The points represent the number of tracks in 0.5 mm overlapping intervals.

the group consists of alpha-particles. We suggest that they are the alpha-particles discussed above and are then to be attributed to the reaction

$$F^{19}+H^1\rightarrow(Ne^{20})\rightarrow O^{16*}+He^4+Q$$
$$O^{16*}\rightarrow O^{16}+h\nu.$$

Using the methods described by Bethe[8] we compute the energy released in the reaction to be $Q=1.74\pm0.10$ Mev so that the excited state of oxygen involved lies 6.2 Mev above the ground state in good agreement with the observed gamma-ray energy. The yield of alpha-particles is apparently smaller than that observed for the gamma-rays. Our results also indicate that the yield above resonance on a thick target is thirty times that of the nonresonance reaction

$$F^{19}+H^1\rightarrow(Ne^{20})\rightarrow O^{16}+He^4+7.95\text{ Mev},$$

in which the total available energy appears as kinetic energy of the heavy particles. These results are based on preliminary measurements at an angle of 90° and may be in error if the alpha-particle distributions are markedly anisotropic.

Further investigations are now in progress to determine more accurately the resonance voltage of the alpha-particles and their distribution in angle with the proton beam.

W. B. McLEAN
R. A. BECKER
W. A. FOWLER
C. C. LAURITSEN

Kellogg Radiation Laboratory,
California Institute of Technology,
Pasadena, California,
March 31, 1939.

[1] Delsasso, Fowler and Lauritsen, Phys. Rev. **51**, 527 (1936).
[2] J. Halpern and H. R. Crane, Phys. Rev. **55**, 260 (1939).
[3] Hafstad, Heydenburg and Tuve, Phys. Rev. **50**, 504 (1936).
[4] W. Bothe and W. Gentner, Zeits. f. Physik. **104**, 685 (1937).
[5] Bernet, Herb and Parkinson, Phys. Rev. **54**, 398 (1938).
[6] Kalckar, Oppenheimer and Serber, Phys. Rev. **52**, 279 (1937); E. R. Gaerttner and H. R. Crane, Phys. Rev. **52**, 582 (1937); see also discussion by Breit in reference 5.
[7] W. E. Burcham and C. L. Smith, Proc. Roy. Soc. **168**, 176 (1938).
[8] H. A. Bethe, Rev. Mod. Phys. **9**, 276 (1937).

The Evolution of Red Giants

It was recently shown by the author and E. Teller[1] that there exist in the mass-luminosity diagram a number of bands parallel to the main sequence in which different thermonuclear reactions of the light elements (D, Li, Be, B) play a predominant role. A star, starting its evolution with an appreciable amount of these elements, should proceed in the direction of decreasing radii and slowly increasing luminosities, and spend a comparatively long time within each of these bands and undergo a more rapid gravitational

contraction during the transitions from one of such regions into another.

It is known, however, that the distribution of red giants in *R-L*-diagram is limited from above by a line on which the pulsating-variables are located, and that between this line and the upper part of the main sequence only very few stars are known.

It seems possible to understand the existence of this limit, and the pulsational instability of the stars located near it, on the basis of the following simple considerations. As was shown by the author,[2] the condition that the nuclear energy production is more important than the energy liberation due to the gravitational contraction can be written in the form: (notations of reference 2)

$$\frac{R}{M} > \frac{\gamma G m}{(n+1.75)\epsilon} \cdot x^{-1}, \qquad (1)$$

where x is the concentration of the element responsible for the energy production. Accepting the concentration of the light elements to be about 10^{-4} (what is necessary to understand the observed lifetimes of the red giants) we obtain:

$$\frac{R}{M} > 3\times 10^{-22} \frac{\text{cm}}{\text{g}}; \quad \frac{[R/R(\odot)]}{[M/M(\odot)]} > 1. \qquad (2)$$

The limiting line, representing this condition in the *R-L*-diagram, has the same general direction as the line of pulsating stars and can be brought into the coincidence with it by a small adjustment of the value of x. Above this limiting line the evolution is always purely gravitational (until the stars gets into the main sequence) and no "lateral" structure, corresponding to different light elements, can be expected. The number of stars observed in this region must be statistically very small because of the short time scale of the gravitational contraction.

These considerations also bring us to the conclusion that the pulsative instability of the stars near the limiting line must be due to the conditions existing during the transition from the state of thermonuclear evolution into the state of purely gravitational contraction.

G. GAMOW

The George Washington University,
Washington, D. C.,
March 24, 1939.

[1] G. Gamow and E. Teller, Phys. Rev. **55**, 654 (1939).
[2] G. Gamow, Phys. Rev. **55**, 718 (1939), formula (6).

The Ionization of Mercury

In view of the importance of the results published recently by Bell and by Nottingham,[1] it seems of interest to mention briefly similar results which have been obtained by a different method. Experiments have been done by the writer[2] with a modified space-charge tube sealed off from the pumps with most of the precautions taken by Nottingham, except that getter was not used. The primary beam was apparently exceptionally homogeneous. It was found that a sharp increase in the ionization always occurred at nearly 10.4 volts. This has been assumed to represent the low energy edge of the main ionization continuum of Hg I, and with this calibration the following ionization processes have been noted.

(1) The current from the space-charge filament decreased continuously as the primary voltage was increased up to five volts, after which ionization set in and varied with the voltage in the manner found by previous workers.[3] It has been found unnecessary to postulate the formation of Hg^- ions to explain the initial drop, because this is due to the effect of the primary electron beam on the space-charge filament.

(2) A rapid rise in the ionization occurred above 10.38 volts with a maximum at 10.81 volts in agreement with the maximum found by Nottingham.

In later runs the ionization between 8.5 volts and 12 volts has been studied by a sensitive method, used previously by R. H. Sloane[4] in analysis of gas discharges, in which there is performed an automatic double differentiation of the current from the space-charge filament relative to the primary voltage. It was found that the second differential curve had first a large positive maximum corresponding to the 10.38-volt inflection and then became almost discontinuously negative to about the same extent, passing through a sharp minimum at 10.50 volts. It is difficult to see how this could be instrumental in origin.

Between 10.5 and 12.0 volts the second differential curve was on the average negative, but rose to positive values at three broad maxima, with intermediate negative minima. The fluctuations correspond to waves in the probability of ionization curve in positions consistent with those found by Nottingham. The points of inflection are those points where the second differential is zero. The maxima in the differential curve had amplitudes about five percent of that representing the 10.38-volt inflection and were about five times as broad. Their character makes it doubtful if they are connected with the onset of new ionization processes, and suggest rather that they correspond to changes in the probability of processes which had commenced at lower voltages.

The second differential curve has also some 20 small sharp peaks between 10.4 and 12.0 volts. Their positions can be correlated with most of the critical potentials previously reported in this region, but the reality of all has not yet been definitely established and the probability of chance coincidences is large.

T. MCFADDEN

Queen's University,
Belfast, Northern Ireland,
March 10, 1939.

[1] M. E. Bell, Phys. Rev. **55**, 201 (1939); W. B. Nottingham, Phys. Rev. **55**, 203 (1939).
[2] Initially in conjunction with G. E. Swindell.
[3] B. L. Snaveley, Phys. Rev. **52**, 174 (1937); F. L. Arnot and M. B. McEwen, Proc. Roy. Soc. **A165**, 133 (1938).
[4] R. H. Sloane, Phil. Mag. **18**, 193 (1934).

Production of Neutrons in Uranium Bombarded by Neutrons*

It is conceivable that the splitting of the uranium nucleus may have associated with it the emission of neutrons. These could either evaporate from highly excited fragments (this process is made more probable by the large neutron excess of the fragments, which lowers the binding energy of the neutrons) or be emitted at the instant of fission. This letter is a preliminary report on experiments

undertaken to ascertain whether, and in what number, neutrons are emitted by uranium subject to neutron bombardment, and also whether the number produced exceeds the total number absorbed by all processes whatever.

A source of neutrons was placed in the center of a spherical bulb, 13 cm in diameter, which was immersed in the center of a large tank of water, 90 cm in diameter and 90 cm high. The activity induced in a rhodium foil placed in the water at various distances from the source was measured with and without uranium oxide inside the bulb. If we denote by $a(r)$ the activity induced at the distance r, the integral $\int a(r)r^2dr$ is proportional to the total number of slow neutrons present in the water. A comparison of the integrals with and without uranium gives, in principle, the possibility of deciding whether or not, with uranium present, there is a net increase in the number of neutrons. This result will be independent of the initial energy of the neutrons since in the water they all become slow neutrons.

With radon+beryllium as a source of neutrons a 6 percent increase in the integral with uranium was found. If this increase were caused only by neutrons emitted by the processes discussed above, it would correspond to a yield of about two neutrons per each neutron captured. Since, however, a radon+beryllium source emits also neutrons with energies of several million electron volts one can also interpret the increase observed, or at least part of it, as due to $(n, 2n)$ processes in which a high energy neutron knocks out a nuclear neutron without being captured.

The experiment was therefore repeated with photoneutrons emitted from a block of beryllium irradiated with the gamma-rays from one gram of radium. Since these neutrons have energies considerably lower than 10^6 ev an $(n, 2n)$ process is highly improbable. Because of the large dimensions and the shape of the beryllium block and the radium source, it was not feasible, however, to secure, in this experiment, the requisite spherical symmetry. The resulting loss of precision makes a comparison of the two integrals inconclusive. Since, however, at large distances from the source the activity with uranium is about 30 percent larger than without uranium, it follows that neutrons of energy larger than that of the photoneutrons are produced in the process. This result is in agreement with the direct observations of Szilard and Zinn; we thank them for informing us of their results prior to publication. Close to the bulb the activity with uranium is, instead, about 60 percent of the activity without uranium. This decrease is due to the absorption of neutrons by uranium; the total absorption cross section of uranium, due to fission and other processes, can be deduced from this result to be of the order of 5×10^{-24} cm^2.

If the volume integral of the differences in the activities with and without uranium be calculated, the contribution from parts in the neighborhood of the bulb is negative and gives the order of magnitude of the number of neutrons absorbed. The contribution of the distant parts of the volume is positive and gives the order of magnitude of the number of neutrons produced. These two contributions are of the same order of magnitude and the present accuracy is inadequate to decide which one is larger. Improved experiments are in progress in order to increase the accuracy.

Roberts, Meyer, and Wang[1] have reported the emission of delayed neutrons subsequent to the bombardment of uranium by neutrons. Such delayed neutrons cannot contribute appreciably to the effects that have been described here. This was ascertained by observing that no activity was induced in the rhodium detector after the removal of the source.

We are indebted to the Association for Scientific Collaboration for the loan of the photoneutron source used in these experiments.

H. L. Anderson
E. Fermi
H. B. Hanstein

Pupin Physics Laboratories,
Columbia University,
New York, New York,
March 16, 1939.

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

[1] Roberts, Meyer, and Wang, Phys. Rev. 55, 510 (1939).

Structure of the Be I Line λ4573A

The beryllium spectrum has been investigated many times with instruments of high resolving power in the hope of finding hyperfine structure. All attempts have hitherto been unsuccessful. All lines, as it has been proved, are single, with two exceptions: Kruger and Wagner[1] found a weak component near the Be II line 4673A and Parker[2] claimed to detect an asymmetrical structure of the Be I line 4573A. In connection with another experiment I reinvestigated the chief Be I lines (4573, 3321 triplet, and 2349A) and Be II lines (3131A doublet) excited in a liquid-air-cooled hollow cathode discharge tube of the type described by Ritschl[3] by means of aluminum-coated Fabry-Perot etalons and Lummer-Gehrcke plates (13 and 20 cm long). All lines were found to be single and no asymmetry could be detected in the line 4573A, although this line was investigated very thoroughly. Fig. 1 is a microphotometric record of an enlargement of a photograph of this line taken with a Fabry-Perot etalon of 27.2 mm separation of plates. The rise of the intensity towards the middle of the interference pattern results from the uneven illumination of the slit of the spectrograph; the anomalous height of the second maximum to the right of the middle results from a small scratch on the photographic plate. The high symmetry of the fringes is evident. If an asymmetry were present it should have been detected, firstly because the resolving power of my instrument is more than three times the resolving power of the Lummer-Gehrcke plate used by Parker in his investigations ($\leqslant$700,000), and secondly because I used lower currents (0.35 amp.) than Parker (0.5 amp.) and as I have remarked, the lowering of current is as important for obtaining small breadths of the line as an intensive cooling. That the lines were very sharp is shown by a comparison of the breadth of the resonance line 2349A in my experiments (0.012A, 0.2 amp.) with that in the experiments of Kruger and

FIG. 1. Microphotometric record of the Be I line λ4573A.

Wagner[1] (0.030A). My supposition is that Parker took a microphotometric record of a fringe of too low an order in the interference pattern where the distribution of intensity in a single fringe is deformed by the very rapid falling of the intensity in the direction to the middle of the pattern. In view of the results of my experiments and of the work of Ballard, Anderson and White[4] on the Be III spectrum the conclusion is that if the beryllium nucleus has a magnetic moment different from zero, it must be very small, much smaller than it was supposed by Parker. The weak component of the Be II line 4673A found by Kruger and Wagner is probably due to a foreign line.

S. MROZOWSKI

Institute of Theoretical Physics,
Joseph Pilsudski University,
Warsaw, Poland,
October 4, 1938.

[1] P. G. Kruger and R. C. Wagner, Phys. Rev. 41, 373 (1932).
[2] A. E. Parker, Phys. Rev. 43, 1035 (1933).
[3] R. Ritschl, Zeits. f. Physik 79, 1 (1932).
[4] S. S. Ballard, O. E. Anderson and H. E. White, Phys. Rev. 47, 256 (1935).

Instantaneous Emission of Fast Neutrons in the Interaction of Slow Neutrons with Uranium*

Recently it became known[1] that uranium can be split by neutrons into two elements of about equal atomic weight. In this fission of uranium the two elements produced have a large neutron excess; moreover they are probably produced in an excited nuclear state. One might therefore expect that these excited fragments instantaneously emit neutrons and that perhaps the number emitted is even larger than one per fission.

One might also expect a delayed emission of neutrons—as was first pointed out by Fermi—if some of the fragments go through one or more beta-transformations before they emit a neutron. Delayed emission of neutrons caused by the action of both slow and fast neutrons on uranium has recently been reported by Roberts, Meyer, and Wang,[2] who find a period of about 12 seconds.

In order to see if there is an instantaneous emission of neutrons from the fission of uranium we have performed the following experiment. We exposed uranium oxide to neutrons which were slowed down by paraffin wax, using as a source of neutrons a block of beryllium from which photoneutrons were liberated by the gamma-rays of radium. A helium-filled ionization chamber connected to a linear amplifier served as a detector for fast neutrons. The ionization pulses of the chamber were observed visually by means of a cathode-ray oscillograph and were recorded by the usual counting arrangement.

Figure 1 shows a diagram of the experimental arrangement. The ionization chamber is covered by a cadmium sheet cap *G* which prevents the thermal neutrons from penetrating to the helium ionization chamber. A cadmium sheet shield *H*, 0.5 mm thick, is used to cover the cylindrical box *E* which contains 2300 g of uranium oxide. The uranium oxide is screened from the thermal neutrons by this shield and can be exposed to them simply by removing the shield.

We observed about 50 pulses per minute from the helium chamber when we exposed the uranium oxide to the thermal neutrons in the absence of the cadmium shield *H*, but obtained only 5 pulses per minute when the uranium was screened from the thermal neutrons by the cadmium shield. The difference of about 45 pulses per minute we have to attribute to fast neutrons emitted from uranium under the action of thermal neutrons. It is reasonable to assume that this emission of fast neutrons is connected with the fission of uranium.

Control experiments were carried out in which uranium was replaced by lead. The effect of the presence and absence of the cadmium shield *H* and the cadmium cap *G* was tested.

In order to estimate the number of fast neutrons emitted per fission under the action of thermal neutrons we used an ionization chamber lined with a thick layer of uranium oxide having an area of 25 cm². This uranium chamber was put in place of the helium chamber without otherwise materially changing the experimental arrangement. Under these conditions the uranium chamber gave about 45 fissions per minute. Assuming the range of the fission fragments to be about 0.005 g per cm² in uranium oxide, the observed 45 fissions per minute should occur in a surface layer, weighing 0.13 g, of the thick uranium oxide lining. Accordingly, about 800,000 fissions per minute should occur in the 2300 g of uranium oxide which was used in our experiment. By taking into account the solid angle, the size of the helium chamber and the pressure used, and by assuming that the "fission neutrons" have an average collision cross section in helium of 3.5×10^{-24} cm² we find the number of neutrons emitted per fission to be about two.

This number is of course only a rough estimate; the main cause of uncertainty is the considerable variation of the cross section of helium with the neutron energy in

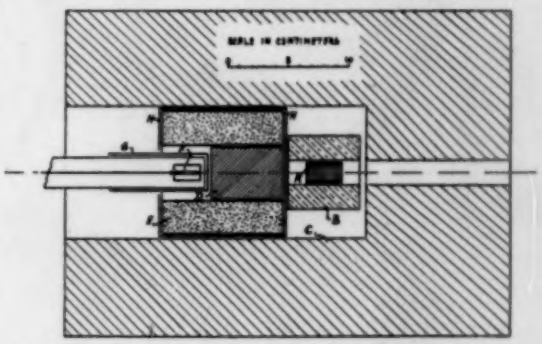


FIG. 1. Arrangement for the observation of the emission of fast neutrons from uranium. *A*, Radium. *B*, Beryllium block. *C*, Paraffin wax. *D*, Lead block. *E*, Box filled with uranium oxide. *F*, Ionization chamber. *G*, Cadmium sheet cap. *H*, Cadmium sheet shield.

the region around one million volts.[3] A hydrogen-filled ionization chamber is now being used in order to obtain a more accurate estimate. It seems to be established, however, that the order of magnitude is one neutron per fission.

Anderson, Fermi and Hanstein have independently, and by a different method, carried out experiments on the neutron emission connected with the fission of uranium. Our observations are consistent with their results, and we wish to thank them for communicating their results to us before publication.

While from our observations we can only say that the time delay involved in this "instantaneous" neutron emission appears to be less than one second, we should expect, for theoretical reasons, this emission to take place within less than 10^{-14} second.

We have also looked for a delayed emission of fast neutrons by performing the following experiment. The uranium oxide was irradiated for some length of time in the arrangement shown in Fig. 1. Then the radium was quickly removed from the beryllium block and the cathode-ray oscillograph screen was watched for a period of 15 seconds for an indication of a delayed emission of fast neutrons. After the radium is removed there is no gamma-ray background to set a lower limit for the observable helium recoil energy; the only slight background remaining is due to electrical fluctuations of the amplifier. In 50 experiments, corresponding to a total observation time of more than 12 minutes, we observed only two pulses which may or may not have been due to a delayed emission of fast neutrons. This is to be compared with the emission of 45 fast neutrons per minute, the number observed while the radium is inside the beryllium block. We conclude that, if slow neutrons falling on uranium cause a delayed emission of neutrons which are sufficiently fast for us to observe, their number must be very much smaller than the number of neutrons which we have observed in the instantaneous emission.

We are indebted to Dr. S. Seely for his assistance in carrying out some of these experiments. We wish to thank the Department of Physics of Columbia University for the hospitality and the facilities extended to us, and also wish to thank the Association for Scientific Collaboration for enabling us to use one gram of radium in these experiments.

LEO SZILARD
WALTER H. ZINN

Pupin Physics Laboratories,
Columbia University,
New York, New York,
March 16, 1939.

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

[1] O. Hahn and F. Strassmann, Naturwiss. **27**, 11 (1939); L. Meitner and R. Frisch, Nature (February, 1939).

[2] R. B. Roberts, R. C. Meyer, and P. Wang, Phys. Rev. **55**, 510 (1939).

[3] H. Staub and W. E. Stephens, Phys. Rev. **55**, 131 (1939).

Proceedings of the Metropolitan Section of the American Physical Society

Meeting of March 24, 1939

THE second meeting of the Metropolitan Section of the American Physical Society for the season 1938–1939 was held at 2:30 P.M. on Friday, March 24, 1939, in the Pupin Physics Laboratories of Columbia University. The following papers were presented:

Present Status of the Analysis of Atomic Spectra. A. G. Shenstone.
Cosmic Rays: The Primary Rays and the Phenomena of Showers. K. K. Darrow.
Cosmic Rays: The Mesotron. Enrico Fermi.

The following officers were elected for the season 1939–1940:

Chairman—I. I. Rabi
Vice Chairman—L. A. Turner
Secretary-Treasurer—W. S. Gorton
Members of the Executive Committee—W. H. Brattain, F. E. Myers

W. S. Gorton,
Secretary-Treasurer